U.S. DEPARTMENT OF COMMERCE
National Technical Information Service

PB-273 195

Survey and Evaluation of Techniques Used in Testing Chemical Substances for Teratogenic Effects

Tracor Jitco, Inc, Rockville, Md

Prepared for

Environmental Protection Agency, Washington, D C Office of Toxic Substances

Oct 77

PB 273 195

REPRODUCED BY
NATIONAL TECHNICAL
INFORMATION SERVICE
U. S. DEPARTMENT OF COMMERCE
SPRINGFIELD, VA. 22161

TECHNICAL REPORT DATA *(Please read Instructions on the reverse before completing)*		
1. REPORT NO. EPA-560/5-77-007	2.	3. RECIPIENT'S ACCESSION NO. PB275195
4. TITLE AND SUBTITLE Survey and Evaluation of Techniques Used in Testing Chemical Substances for Teratogenic Effects		5. REPORT DATE October 1977 6. PERFORMING ORGANIZATION CODE
7. AUTHOR(S)		8. PERFORMING ORGANIZATION REPORT NO.
9. PERFORMING ORGANIZATION NAME AND ADDRESS Tracor Jitco, Inc. 1776 East Jefferson Street Rockville, Maryland 20852		10. PROGRAM ELEMENT NO. 11. CONTRACT/GRANT NO. 68-01-2204, Task #29
12. SPONSORING AGENCY NAME AND ADDRESS Office of Toxic Substances U.S. Environmental Protection Agency Washington, D.C. 20460		13. TYPE OF REPORT AND PERIOD COVERED Final 14. SPONSORING AGENCY CODE 560/12
15. SUPPLEMENTARY NOTES		

16. ABSTRACT

This survey summarizes and evaluates the methods currently used or potentially useful for testing chemicals for teratogenic effects.

17. KEY WORDS AND DOCUMENT ANALYSIS		
a. DESCRIPTORS	b. IDENTIFIERS/OPEN ENDED TERMS	c. COSATI Field/Group
Teratology Testing methods Teratogenic chemicals		
18. DISTRIBUTION STATEMENT	19. SECURITY CLASS *(This Report)*	21. NO. OF PAGES 102
	20. SECURITY CLASS *(This page)*	22. PRICE PC A06 / MF A01

EPA Form 2220-1 (9-73)

EPA-560/5-77-007

Survey and Evaluation of Techniques Used in Testing Chemical Substances for Teratogenic Effects

Final Report

October 1977

Contract 68-01-2204, Task 29

Office of Toxic Substances
Environmental Protection Agency
Washington, D.C. 20460

Lois Jacob
Project Officer

TABLE OF CONTENTS

		Page
I.	Introduction - Objectives and Scope	1
	A. Objectives	1
	B. Scope	2
	C. Difficulties of Teratogenicity Testing	3
II.	Problems of Design	4
	A. Extrapolation from Animals to Man	4
	B. Designing a Prospective Animal Testing for Teratology	6
	C. Selection of Control Parameters	8
	D. Knowledge of the Chemical to be Tested	10
	E. Expected Types of Malformations	11
III.	Currently Available Methodologies and Their Evaluation	12
	A. In Vivo Methods	13
	1. Species and Strain of Test Animal	13
	2. Administration of Chemical	19
	3. The Test Environment	36
	4. Observations	38
	5. Interpretation of Results	45
	B. In Vitro Systems	50
IV.	Currently Used Methodologies	53
V.	Surveillance	57
VI.	Economics	61
VII.	List of Known Teratogens	64
VIII.	Recommendations	65
	A. Screening Chemicals Already in the Environment for Teratogenic Effects	65
	B. Prospective Screening of Chemicals Not Yet in the Environment for Potential Teratogenicity	69
IX.	Possible Future Methodologies	74
	A. Postnatal Evaluations	74
	B. Possible Shortcuts in Current Procedures	74
	C. Additional Species of Test Animals	75
	D. Improved Monitoring of Human Populations	76
X.	Concluding Remarks	77

LIST OF ILLUSTRATIONS

Figure #	Page #	
1	43	Cleared, Alizarin - stained monkey fetus removed by hysterotomy from an untreated rhesus female on day 100 of pregnancy.
2	44	Radiograph of a monkey fetus removed by hysterotomy from an untreated rhesus female on day 100 of pregnancy.

LIST OF TABLES

Table #	Page #	Title
1	5	Gestation Stages
2	14	Some Animals Used or Available for Teratogenicity Testing
3	15	Animals Recommended or Required by Various Contries for Drug Teratogenicity Testing
4	15	A New Concept on Teratogenicity Testing Based on Multilevel Tests in Different Types of Animals
5	17	Comparative Thalidomide Teratogenicity
6	17-18	Thalidomide Teratogensis in Primates
7	18	Percent Malformations Produced in Four Inbred Strains of Mice by Galactoflavin
8	19	Range of Spontaneous Malformations in Animals
9	20	Teratogenicity Toxicity Relationships for Some Drugs in the Rat
10	27	Episodes on the Reproductive Cycle of Mammals
11	28	Gestation Times for Some Laboratory Animals and Man
12	29	Critical Periods of Organogensis in Animals
13	30	Some Chemical Agents Known to Influence Rates of Metabolic Degradation of Themselves and/or Other Repeated Dosages.
14	31	Ways in Which Repeated Treatment Prior to the Peak Susceptible Period of the Embryo May Produce Misleading Results
15	32	Numbers of Pregnant Rats Treated at Consecutive 3-day Periods
16	32	Types of Tests for Developmental Abnormality Depending on Duration of Treatment
17	42	Weights and Measures on 15 Control Monkey Fetuses Removed by Hysterotomy on Gestation Days 99-101 From Untreated Rhesus Females
18	52	Some In Vitro Systems that have been Employed in Teratology
19	62	Typical Costs for Some Carcinogenesis Bioassay Functions Determined by Tracor Jitco, Inc. for the Cancer Institute in 1976.
20	63	Laboratory Procurement and Maintenance Costs Determined by the American Association of Medical Colleges for Year Ending June 30, 1973.

I. INTRODUCTION - OBJECTIVES AND SCOPE

Adverse effects on human reproduction during early gestation are frequent. According to the March of Dimes National Foundation, of the three million births in the United States each year, about 200,000 (6.7%) of all newborns have birth defects and more than 560,000 (18.7%) of all pregnancies terminate in spontaneous abortions, stillbirths or miscarriages caused by maldevelopment (101).

Concern over the potentially harmful effects of drugs taken during pregnancy is relatively recent, dating back to the "thalidomide tragedy" which occurred in the 1960s following introduction of this sedative-hypnotic compound, into the market in 1956. Its use during early pregnancy was shown to result in phocomelia. Since then other compounds such as androgens, progestrogens, aminopterin, methotrexate, and antithyroid drugs have been positively identified as human teratogens. Compounds regarded as "suspect teratogens" include anticonvulsants, neurotropic, anoexogenics, oral hypoglycemics and alkalylating agents, while those regarded as "possible teratogens" include aspirin, antibiotics, quinine, barbiturates, etc. (128).

Compounding the problem of women of child-bearing age exposed to drugs during pregnancy is the problem of occupational exposure to chemicals in various laboratories and industrial work places. It is estimated that 2,000 or more new chemicals are produced each year, and at least 200 of these are released into the environment in amounts that could cause developmental disturbances (108). For this reason the testing of chemicals used in agriculture, food processing and production and industry is now considered mandatory by most regulatory agencies. The safety assessments of these chemicals are based on experimental data obtained for several species tested by employing various teratologic approaches.

A. Objectives

This survey is intended to summarize and evaluate the methods currently used or potentially useful for testing chemicals for teratogenic effects. The scope of the survey is limited as described hereunder.

B. Scope

A teratogen, as conventionally understood, is an agent that causes structural defects or anomalies of prenatal development that are present, but not necessarily detectable, at birth. Such defects are often called congenital malformations. This concept may be considered narrow because it neglects other facets of embryotoxicity such as prenatal death, overall growth retardation, and functional deficits. Usually such deficits, when caused by a teratogen, become detectable long after birth and only then may be traced to a prenatal origin.

The environmentally-induced developmental defects covered in this report include intrauterine death, fetal growth retardation, structural defects, and early-appearing functional deficits. A limitation imposed on this report is that the defects shall have originated after implantation, which excludes defects that are hereditary, whether genic or chromosomal. A limitation on the coverage of functional deficits is that both behavioral studies and observations later than the immediate postnatal period are specifically excluded.

Many environmental agents have been found to cause developmental defects in experimental animals, including many that are outside the scope of this report (e.g., radiation, noise, temperature extremes). The report considers only methods used for testing whether chemical substances introduced into the environment cause developmental defects when females have been exposed to them during pregnancy. Defects caused by chemical deficiencies are included only if they result from positive introduction of other chemical(s) into the environment, and the test methodology involves a test for the deficiency.

Studies to evaluate food chemicals and therapeutic drugs for teratogenicity were not analyzed in depth unless there was some unique and appropriate aspect of the experimental design.

This survey is evaluative and based on a literature search and by information from consultants; it covers all methods that have been used or have been recommended for use during the past 50 years. It also covers some methods which the evaluation suggested might be considered or reconsidered in the future. It does not cover every method that has ever been used for teratogenicity testing.

The methods surveyed are those particularly directed to (a) testing chemicals for teratogenicity and (b) surveillance of human populations for teratologic manifestations related to environmental chemicals. The chemical may enter the proximate or ambient environment by whatever means; the report is not concerned with how or in what amounts it enters the environment. However, a central question is, How may pregnant human females be exposed? and tests for teratogenicity are generally designed according to the likely route(s) of human exposures. Test findings are considered only to the extent that they are criteria for the evaluation of methods.

C. Difficulties of Teratogenicity Testing

The ultimate question that a test method is designed to answer is, Could this chemical harm unborn children? Direct experimentation being impossible, the question becomes, How can an answer to this question be inferred? or, What is the probability of adverse developmental effects? It is clear that no available prospective methodology achieves the full objectives. At best, available methods estimate degrees of probability that a chemical may be teratogenic for human beings. Even retrospective surveillance methods often reach equivocal conclusions. This report therefore addresses questions such as What animal testing methods are currently used, what available methods come nearest to providing a valid answer, what are the limits of confidence of the best available methods, and what other methods may be devised?

II. PROBLEMS OF DESIGN

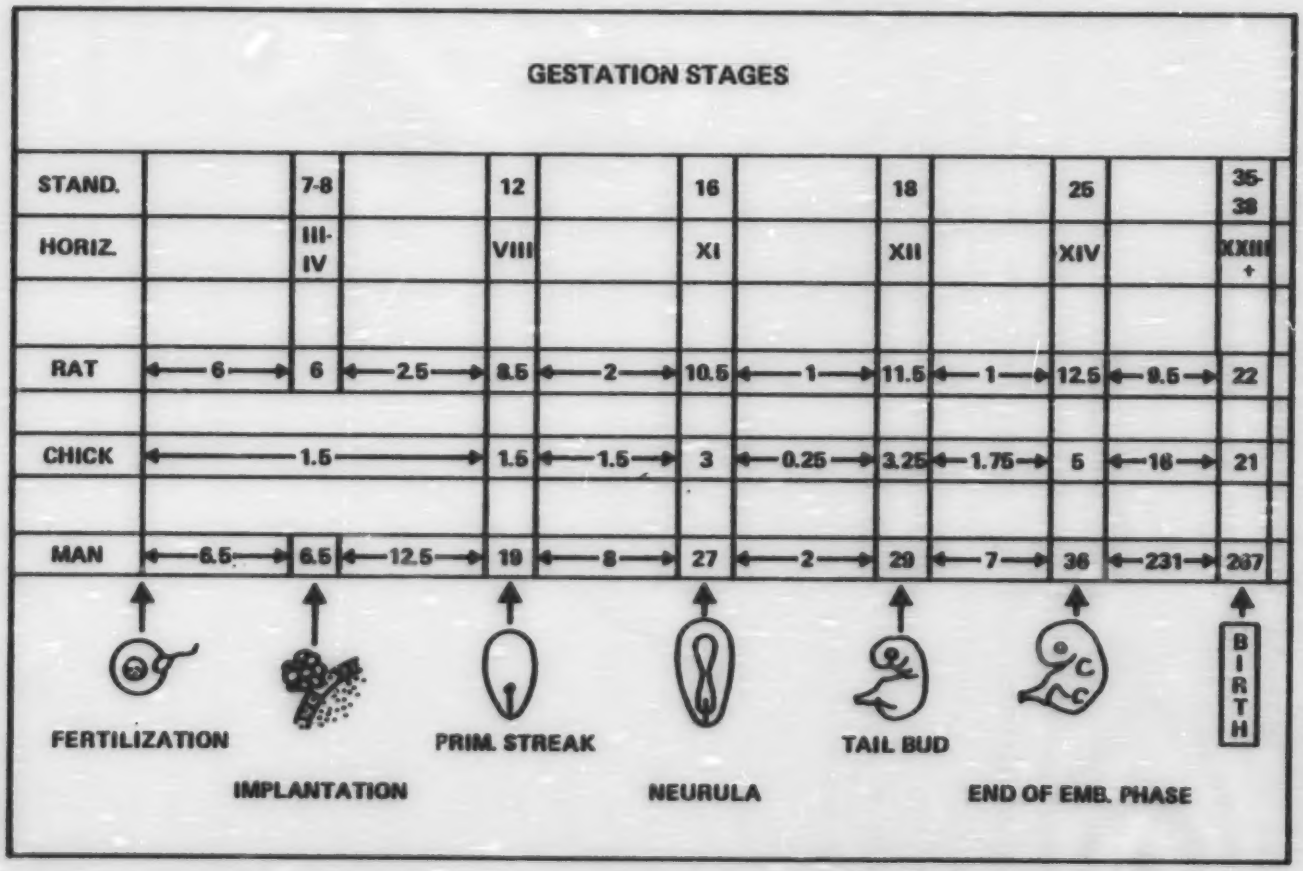
A. Extrapolation from Animals to Man

The thorniest problem undoubtedly is that of extrapolability to man. No single factor - whether related to duration and level of exposure, response, timing, or environment - can be transferred directly with acceptable confidence. Retrospective surveillance of human populations is the only direct approach that is at present fully acceptable and feasible. Even then, surveillance is often incomplete and unrewarding in practice. The thalidomide disaster demonstrated that nonspecific retrospective surveillance was inadequate. Prospective animal tests must therefore be designed to best infer and assess any potential teratogenicity of agents for man.

In the absence of knowledge of the many possible factors (metabolic, embryological, and other) that may underlie susceptibility of human embryos to a given chemical, species and strains of test animal cannot be chosen rationally. These choices must normally be arbitrary. Also, the variables of the test design must be approached empirically. For example, the nature of the test environment, amounts of exposure to the chemical, timing of exposures, responses to be looked for, and frequency of such responses in the test animal population must to some extent be decided for each suspect chemical individually. Final, conclusive evidence of teratogenicity for man must be sought in the offspring of women exposed during pregnancy after a chemical has entered their environment.

There are many reasons for this uncertainty. An important one is that the time course of intrauterine development varies among species without regard to lifespan or gestation period. For instance, the preimplantation period (during which cells are undifferentiated, so there is no basis for teratogenesis) varies from 4.5 days in hamsters to 6.5 days in man to 10 days in sheep while the total gestation period varies from 16 days in hamsters to 278 days in man (69). Organogenesis occupies the first 6 weeks of human development, but overlapping and telescoping of the steps of development vary among species, so that comparisons are usually difficult (Table 1).

Table 1 (98)



Abbreviations: Horizons shown in Roman numerals, represent a series of 23 stages in the development of the human embryo (139); Standard, shown by Arabic numerals, represent a series of 36 stages in the development of the human embryo (169).

The chronology of structural developments in the central nervous system (CNS), especially at the histological level, varies greatly among species, and the final steps (postnatal) probably occur later in man than in most test animals. The CNS is a prime target for teratogenesis. But there is little comparative knowledge about development, as was stated in 1967 by the WHO Scientific Group (173).

Other differences between man and many test animals that may be important in teratology include: placental differences; single implantation in humans versus multiple implantations in many test animals (relevant to resorption of embryos and fetuses, and to abortions); and differences in endocrinology, metabolism, pharmacology, pharmacokinetics, and nutrition. Exposed human populations are usually larger and more randomly bred than are groups of test animals, and so they are genetically more heterogeneous.

Finally when a compound has been demonstrated to be teratogenic in animals, it must be determined whether the mechanism is specific to the animal under investigation or whether it is also relevant to man. Information about the mechanisms involved in maternal and fetal toxicity and teratogenesis could allow some extrapolations to be made from animal models to man (144).

B. Designing a Prospective Animal Test for Teratogenicity

Experimental conditions and test designs contain many variables that can influence the interpretation of results. These must be recognized and then carefully standardized or regulated.

Experimental variables include: all aspects of husbandry such as quality of feed, bedding, drinking water; temperature, humidity, barometric pressure, amount and periodicity of light and noise; cage size, material, type of racks; group size unless animals are caged individually; state of health and standard of laboratory care; species and strain of test animals; parity and time of mating of test females; exposures of test females to the chemical (dose range, dose timing, route of exposure, duration of treatment) decided largely by the nature and purpose of the test chemical; and treatment of concurrent control animals.

An important factor that must be considered in the design of an experimental model is the pharmacological activity between the maternal and fetal compartments (48, 84). In general, factors that alter the level of free drug or its metabolite in the maternal compartment are capable of altering the level of free drug in the fetus as well. The degree of transfer of drugs from the placenta is directly proportional to the free drug concentration in the maternal plasma. Since the fetal system has no other means of drug excretion except through redistribution to the maternal compartment, any metabolism of the drug within the fetus or the placenta would result in containing that drug within the fetus and lead to drug-induced teratogenesis (55).

The design of concurrent control treatments should never be approached in a routine manner (see Section II, 3). The valid observation of results can demand skilled insight at the design stage. What to look for and when to look for it should be decided in detail according to the priorities of each particular test. For example, a choice of methods or of timing may be available for recovery of offspring from treated females, for examinations of fetuses, and for postnatal examinations of offspring not killed at birth. On the other hand, all teratogenicity tests require meticulous record-keeping to be kept throughout, necropsy of any animal dying spontaneously, and reporting of all dead and resorbed conceptuses.

When observations are complete, the ways in which the data should be analyzed and interpreted will have been decided as part of the test design. The methods for determining statistical significance and the uses to be made of significant data should be chosen according to the possible sources of error, as foreseen, and the anticipated limits of confidence.

Therefore, only well-qualified experts should undertake the design of teratogenicity tests. Such experts will know the extent to which they can control each of the experimental variables in the actual circumstances. They will make informed choices when certainty is impossible and will be aware of controversies.

For example, the approach to dosage with potential teratogens has been much debated. A program of single doses to test specific sensitivity of every tissue at the time of its rapid proliferation has been recognized as ideal but far too costly (115). However, multiple doses have been stated to produce, on the one hand, cumulative effects (115), or on the other, fewer malformations owing to development of metabolic tolerance (166). To escape this difficulty, dose levels weighted for known tissue sensitivities ("equivalent" doses) have been advocated (115). Although the "threshold" concept of dose-responsiveness is not accepted universally, it has been recommended for planning purposes (166). The concept of a "teratogenic ratio" (toxicity of a chemical to the test female versus its toxicity to the fetus) has been suggested as a possible predictor of teratogenicity for man, but the usefulness of this ratio is disputed (122). Thus, the well-informed investigator has a wide choice of approaches to dosage available at the design stage.

Studies intended to discern relationships between embryonic exposure and teratogenic response should incorporate the following: (1) ability to distinguish between the parent compound and its metabolite, (2) equality of dosages, (3) similarity of treatment schedule, (4) use of comparable developmental stages and (5) determination of embryonic levels before and after attainment of peak levels. Other essentials to be considered include the refinement of techniques to measure levels of teratogens to minimize errors from embryonic loss or leaching and the use of large sampling units to allow for variances within and among experimental litters.

C. Selection of Control Parameters

That controls should be appropriate goes without saying. However, the difficulty lies in defining what is appropriate, since one can never ascertain all of the relevant variables.

Concurrent controls should, first, be contemporaneous with the test procedures. They should duplicate the test conditions in every conceivable way (except one - the test chemical): namely, administration of the test vehicle, all aspects of animal husbandry, health, selection of the test group animals.

In general, however, background knowledge of so-called spontaneous malformations and the range of variants (skeletal and visceral) characteristic of the selected strain of test animal may be of great value. An increase in an infrequent type of defect may not occur in the small groups used in any one test, or if it does occur, it may not be correctly interpretable in that context alone.

On the other hand, information from "positive controls" (responses of test animals to known teratogens) is regarded by many investigators as either irrelevant to the substance being screened, or simply redundant. If animals are available for this purpose, it may be more meaningful to add them to the treatment groups.

Many pitfalls in the design of controls have been discovered. Cases tend to be individual but some examples of the sort of problem will perhaps illustrate why it is necessary to think carefully in advance.

In a study of a pesticide, intermittent exposure was designed. Later, the aerosol was found to have been adsorbed onto room surfaces from which it had evaporated slowly and continuously, thus eliminating the intended no-treatment periods (149).

Intermittent exposure can also be rendered continuous by an affinity of the substance for the subject's own tissues. "Paraquat," a herbicide with delayed lethality, was found to be retained by lung-surface tissues, while systemic uptake depended on how rapidly the tissues could elaborate ATP; thus after intermittent exposure, the actual uptake was continuous (6).

Another example of brief or intermittent exposure producing a continuous challenge and resulting in delayed fetotoxicity involves the azo dyes, such as trypan blue. These dyes become concentrated in macropnages and are released slowly; after brief exposure of a pregnant test female, the dye persists in her plasma, so that the conceptuses are exposed for several days after the brief maternal exposure. Strontium-90 is another chemical that is released slowly and continuously after brief exposures.

D. Knowledge of the Chemical to be Tested

It is important to know in advance how the chemical is metabolized, as Cahen has pointed out (13). Often there are differences between species or strains of animals in rates and pathways of metabolism and elimination, which might affect the exposure of a conceptus to teratogenic effects of a chemical or its metabolites. Usually the manufacturer or principal user of a chemical is the major or only source of such knowledge. Pertinent information includes, at the least, purity and stability of the chemical, its cumulative toxicity, physiologic (tissue) storage capacity, rate of transport through maternal blood and storage and metabolism by the developing fetus.

Until a few years ago, placental membranes were regarded as impermeable to drugs and chemicals. It has now been shown that most drugs readily traverse the placenta, with permeability modulated by lipid solubility, molecular weights, age of placenta, etc. (28).

Measurements of plasma half-life, protein binding metabolism and tissue distribution in adult animals alone cannot provide information on the differences in susceptibility among offspring of a litter. A combined analysis of these parameters, along with placental transfer kinetics and embryonic localization of the test chemical and its metabolites will enable one to relate both the dose administered to the mother and the dose reaching the fetus to the observed teratological effect.

It is also necessary to consider effects that originate from metabolic interactions within the maternal-fetal system. These effects may be due to microsomal activity of the liver, placenta and/or fetus, which could be inhibited or stimulated. For example, thalidomide has been regarded as teratogenic mainly during early-stage organogenesis. However, free glutamic acid has been found to produce fetal abnormalities during late-stage CNS development (150), and free glutamic acid is a major metabolite of thalidomide. Although this metabolite does not seem to be responsible for the early-stage effects of thalidomide, the question of whether thalidomide-derived glutamic acid could cause late-stage CNS defects remains unanswered.

E. Expected Types of Malformations

The investigator who looks only for particular types of malformation may be misled. To some extent, the time of exposure can influence the type of malformation when a chemical is more toxic to the conceptus than to the mother (13). More importantly, the possible effects of a previously untested chemical cannot always be entirely foreseen. Therefore, when a chemical is to be screened for potential teratogenicity, the observations should be designed to detect any type of developmental anomaly.

III. CURRENTLY AVAILABLE METHODOLOGIES AND THEIR EVALUATION

A. In Vivo Methods

1. Species and Strain of Test Animals

Animals used for teratogenicity studies, according to a 20-year literature search, are listed in Table 2. Species recommended or required by various countries for testing teratogenicity of drugs are listed in Table 3. Rats, mice, and rabbits have been used for the most part.

Wilson (163, 165) has recommended a multilevel approach, using a rodent (rat, mouse, hamster, or rabbit) at the first level, a carnivore (dog, cat, or ferret) or ungulate (pig or sheep) at the second and third levels, and a non-human primate (monkey or baboon) at the fourth level, as shown in Table 4. Properly applied, this approach takes into account such factors as numbers of individuals expected to be exposed, levels of exposure, and cost-risk-benefits.

The multilevel approach acknowledges that there is no ideal animal species for testing for potential teratogenicity in man. At present, there is little experience recorded using this approach.

The ideal species would resemble man as to absorption, metabolism, and elimination of chemical substances, placental transfer of those substances and their metabolites, and course of development of the conceptus (165). Because these and other attributes are either known or believed to influence fetal responses to environmental factors, much research has concentrated on degrees of resemblance for one or another of these attributes. Other research has looked for similar responses to particular substances in man and various animal species.

Neither of these directions of search has been outstandingly productive. Wilson points out (165) that similarity of response between any species and man varies from one substance to another. Kalter notes that reported similarities of attributes tend to be incomplete and poorly

substantiated, that the reasons for the differences must be better understood before a test species can be chosen on a fully rational basis, and that meanwhile it might be more helpful to compare various species with one another than with man.

For example, placental similarity is recognized as desirable (163). On the other hand, many major structural malformations arise before a true placenta is established in any species, even man (115). Kalter notes that much teratogenicity occurs before placental differences become important. Thus, Tuchmann-Duplessis (152) urged the use of small rodents for screening, and use of primates only when specially indicated. One such indication might be a suspicion of late-stage teratogenicity for man.

Responses to particular substances vary greatly among species. This is shown by Kalter for thalidomide in Table 5, and he also emphasizes the genetic variations among strains of a species which influence the teratogenic response. Table 6 shows the responses of various primates to thalidomide, and Table 7 shows the responses of four strains of mouse to galactoflavin.

The reaction to a specific teratogen may also be the result of differences in the rate of metabolism and the qualitative differences in metabolic pathways among species. For example, the role of the fetal liver in the rat, guinea pig, rabbit and swine, with reference to drug metabolism is negligible or absent, whereas in humans the fetal hepatic site is very active. Although the plasma half-life of thalidomide is the same in the rabbit (susceptible species) and the rat (nonsusceptible species), the drug is more readily absorbed and its metabolites more slowly excreted in the rabbit than in the rat (128). Even when two species metabolize a drug at the same rate, the metabolic products may be different and cause different teratogenic responses in the two species (e.g, imipramine) (153).

A number of different animal species have been used in an attempt to determine the most satisfactory model for teratological research for predicting the hazard to man. However, none have been totally adequate because of high spontaneous malformation rate (mouse) (Table 8), low

Table 2
Some Animals Used or Available for Teratogenicity Testing

Animals	Refererceas
Invertebrates:	
Echinoderms:	
Sea urchins	139
Sand dollars	75, 76, 169
Insects:	
Fruit fly	15, 45
Vertebrates:	
Fish:	11, 67, 134
Amphibians:	
Frogs	50, 160, 174
Toads	91, 146, 147, 174
Salamanders	43, 61
Reptiles:	
Turtles	96
Birds (embryonated eggs):	
Chicken	5, 13, 16, 34, 115, 164, 169
Quail	22
Pheasant	135
Mammals:	
Marsupials:	
Opossum	69, 169
Edentates:	
Armadillo	5, 13, 33, 69, 115, 169
Rodents:	
Mouse	5, 14, 38, 69, 122, 165, 173
Rat	5, 18, 69, 122, 153, 163, 165, 173
Guinea pig	5, 13, 34, 60, 69, 169
Hamster	5, 13, 19, 31, 69, 163, 165
Lagomorphs:	
Rabbit	5, 14, 69, 122, 163, 165, 173
Hare	165
Carnivores:	
Mink	5, 27
Ferret	5, 99, 139, 163, 165,
Cat	5, 69, 81, 115, 153, 163, 173
Dog	5, 69, 115, 153, 163, 165, 173
Ungulates:	
Pig (regular, miniature)	5, 69, 115, 124, 163, 165, 173
Goat	69
Sheep	5, 69, 110, 115, 163
Cattle	69
Nonhuman primates:	
Prosimian:	
Bushbaby	5, 110, 153, 173
Simian:	
Monkeys (macaques, marmoset)	5, 69, 153, 163, 165, 169, 173
Apes (baboon)	5, 58, 69, 163, 165

Table 3
Animals Recommended or Required by Various Countries for Drug Teratogenicity Testing

Country	Animal	Reference
USA (FDA)	At least 2 species. Mouse, rat, rabbit, acceptable.	78, 153
Canada	At least 2 species (one a non-rodent). Mouse, rat, hamster, rabbit.	5
Great Britain	Rat and rabbit.	122, 153
France	Three species. Rat, mouse and rabbit.	153
Japan	Two species. Rat, mouse, rabbit used.	115
Sweden	Two species.	34
Switzerland	Rat, mouse, rabbit.	115

Table 4
A New Concept in Teratogenicity Testing Based on Multilevel Tests in Different Types of Animals (165)

Order of test[a]	Purpose	Suitable species	No. of Pregnant Animals
First level	Find embryotoxic dose range	Rat, mouse, hamster or rabbit	130-150
Second level	Confirm or adjust above	A carnivore or an ungulate	40-60
Third level	Only if second level results are equivocal	Alternate to that used in second level	40-60
Fourth level	Only if use in human pregnancy needed or likely	Macaque monkey or baboon	40-50

[a]Tests would terminate at second or third levels in most instances.

Source: Environment and Birth Defects by J.G. Wilson, reprinted with permission of the Academic Press. Year of first publication, 1973.

sensitivity to teratogens (rat), absence of pure strains (rabbit) high cost, and limited availability and breeding capabilities (dog, monkey). Although the monkey was an excellent predictive teratological model for thalidomide the use of this animal for testing other drugs (e.g aspirin, aminopterin) gave disappointing results and the response differed greatly from that of man.

According to a WHO Scientific Group (173), all substances known to be teratogenic for man have shown some teratogenicity in the rat, mouse, or rabbit; yet "there is no absolute assurance" that negative tests in these species will predict absence of teratogenicity for man, or that substances that are teratogenic at high doses in these species will be teratogenic for man at lower doses.

Therefore neither current knowledge of physiology nor teratological response patterns give secure grounds for preferring any one species of mammal for predictive screening for potential teratogenicity for man. Nor are there grounds for believing that two or more species would be more predictive than two or more sufficiently divergent strains of one species. Under these circumstances other considerations become important, such as aspects of husbandry, reproduction rates to generate "significant" numbers of offspring, and cost-risk-benefit factors.

Table 5
Comparative Thalidomide Teratogenicity (69)

Species	(mg/kg/day) Smallest Dose Producing Defects	Largest Dose Producing No Defects
Man	0.5-1.0	?
Baboon	5	--
Monkey, Cynomolgus	10	--
Rabbit	30	50
Mouse	31	4000
Rat	50	4000
Armadillo	100	--
Dog	100	200
Hamster	350	8000
Cat	---	500

Source: Reprinted from Teratology of the Central Nervous System by H. Kalter, reprinted with permission of the University of Chicago Press. Year of first publication, 1968.

Table 6
Thalidomide Teratogenesis in Primates (128)

Species	Teratogenic dose mg/kg oral route	Gestation days treated	Defects
Man	0.5-1	20-36	Limbs (80%); ear (20%)
Cynomologous Monkey	10	22-32	Limbs (67%); teratomas (33%)
Baboon	5	18-44	Limbs and tail (40%)
Rhesus Monkey	12-19	24-26; 27 or 30	Limbs (100%)

Table 6 (Cont'd)
Thalidomide Teratogenesis in Primates (128)

Species	Teratogenic dose mg/kg oral route	Gestation days treated	Defects
Bushbaby	20	16-30	
Japanese Monkey	20	24-26	Limbs (100%); tail (17%); CNS (17%)
Stump-tailed Monkey	5-10	24-30	Limbs (100%); tail (20%)
Marmoset	45	25-35	Limbs, ear and jaw (100%)
Bonnet Monkey	5-30	24-29 or 41-44	Limbs (45%); Visceral (52%)

Source: Reprinted from Drugs as Teratogens, by J.L. Schardein, CRC Press, 1976.

Table 7
Percent Malformations Produced in Four Inbred Strains of Mice by Galactoflavin (69)

Strain	Cleft Palate	Skeletal Defects	Brain Defects	Atresia of Esophagus
A/J	3	38	55	78
DBA/1	41	92	44	83
129	8	84	9	30
C57BL/6	13	0	5	2

Source: Reprinted from Teratology of the Central Nervous System by H. Kalter, reprinted with permission of the University of Chicago Press. Year of first publication, 1968.

Table 8

Range of Spontaneous Malformations in Animals

Species	% Incidence Reported
Mouse	0.4-18.6
Rat	0.02-0.85
Rabbit	0.74-6.3
Dog (Beagle)	0.17-1.9
Monkey	0.11-16.8
Cat	1.24
Sheep	0.2-0.3
Pig	0.6-9.8

Source: Data abstracted from reference 128 (Tables 2-10).

2. Administration of Chemical

Dosage levels, routes of administration, durations of treatment, vehicles and other control aspects are all considered together in designing a test. For instance, in screening chemicals for potential teratogenicity, there is debate about whether to give well-timed large doses to uncover any propensity to cause malformations under test conditions, or to give chronic treatment with small doses that approximate the usual conditions of human exposure (39, 165). Both sorts of information may be required, and since environmental "accidents" can cause occasional acute exposures, most initial screening seeks to uncover any potential teratogenicity. In this subsection of the report, the parameters of dosage are examined one by one.

a. Dosage level. Four major overlapping effects of dose level have been described under test conditions: lethal to the mother, lethal to the conceptus, teratogenic, and no-effect ((165), and see Figure 1). The teratogenic range is looked for empirically (69), either from above (13), or from below.

A commonly preferred reference point is the maternal LD_{50} (39, 165), normally established by acute toxicity tests before a chemical is screened for potential teratogenicity. However, since in most instances there is great disparity between the maternal toxic dose and the teratogenic dose (Table 9) (18), the embryo usually has a higher susceptibility than the adult. This is due to a particular vulnerability of certain embryonic cells which are absent in the adult (153). Drugs with the greatest difference between the two values present the greatest hazard.

Table 9

Teratogenicity Toxicity Relationships for Some Drugs in the Rat (18)

Drug	Approximate LD_{50} (IP)[a]	Minimal Teratogenic Dose (IP)	Ratio
Cyclophosphamide	40	7	0.17
Uracil mustard	1.25	0.3	0.24
Chlorambucil	24	8	0.33
Nitrogen mustard	2.0	0.5	0.25
Triethylene melamine	1.25	0.3	0.24
Triethylene thiophosphoramide	8.0	3.0	0.37
Busulfan	60	18	0.3
6-Mercaptopurine riboside	2000	15	0.007
6-Mercaptopurine-3N-oxide	200	50	0.25
6-hydroxylamino purine	800	400	0.5
5-Fluoro-2'-deoxyuridine	1500	0.15	0.0007
5-Bromo-2'-deoxyuridine	1500	100	0.06
Cytosine arabinoside	3000	20	0.006
Aminonicotinamide	15	5	0.33
Hadacidin	5000	3500	0.7
Procarbazine	400	25	0.06

a) Data abstracted from Reference 18.

b) i.p., intraperitoneal

When the LD_{50} data refer to another strain of animal or to different test conditions, either the appropriate LD_{50} should be ascertained, or the figures used should be properly defined. The dosage plan is expressed as fractions of the LD_{50}; there is usually a threshold for lethality, and the range LD_1 through LD_{99} is relatively narrow, so that (for example) the LD_{25} will hardly ever be one-half of the LD_{50}. Therefore, percentages of the LD other than LD_{50} are not used. However, the MTD (median toxic dose) and ED_{50} data are used by some workers as reference points (173).

Dose measurements for teratogenicity testing are usually based on body weight (mg/kg) (173). Little discussion was found on whether this is really the best basis for evaluating exposures to environmental chemicals that are not ingested. For example, specific skin surface area (cm^2/kg body weight) might conceivably be a better basis of measurement in view of the large differences in this ratio between most test animals and man (41, 120). But data have not been expressed on this basis in most studies reviewed (142).

The LD_{50} is obtained as an acute toxicity measurement (see above). Usually a test of chronic LD_{50} would give different data (lower or higher) but the test would contain more variables, and the amount of time that would be required to produce the chronic effects would vary from chemical to chemical. Chronic toxicity measurement might be more appropriate for some environmental chemicals, but not for others (71, 167).

In any case, the basis of measurement is only a starting point, and the dose level actually found to reveal the embryotoxic threshold is used for setting of safe tolerance levels for the pregnant animal (71, 167).

To screen a chemical for potential teratogenicity, Wilson (165) recommends starting with three dose levels: 0.5, 0.25, and 0.125 of the maternal LD_{50}. If these do not reveal the embryotoxic range, other fractions (binary system, usually smaller) are added. Three levels of dosage, determined empirically or by some arbitrary formula, are recommended or required in Canada (5), Great Britain (122), France (153), and by the WHO (173). Two levels are required in Sweden (34). The United States FDA requires two levels

for single-generation tests for teratogenicity of drugs, and three for multi-generation studies (35, 78). Multigeneration studies are not relevant to the present survey and may obscure the assessment of post-implantation teratogenicity.

b. Route of administration. An accepted principle is that the test substance should be administered to the animals by the route that most closely resembles that of the anticipated human exposure (5, 71, 165, 167). There may be exceptions when differences of physiology affect distribution of the test substance.

Thus test substances are given by mouth, by inhalation, by intranasal instillation, by skin application, etc. (5). Occasionally, to facilitate exposure of the test animal, the principle is ignored, and the substance is injected parenterally. In some cases the route of administration has influenced the degree of teratogenic response (16).

In all cases the route will affect some aspect of bioavailability: for instance, its rate, the available proportion of the dose, the available chemical form of the test substance, the relative systemic exposures of the body tissues including the conceptus, or the durations of such exposures. In practice, the principle of resemblance is often breached to make sure that the test animals receive measured doses. Another breach of the principle is administration by a single route when human exposures can be anticipated by more than one route. Every such addition of certainty to the test situation detracts from its interpretability in terms of anticipated human exposures and of their potential teratogenic effects.

These difficulties involve each of the usual routes of administration of test substances:

(1) Oral doses are given preferably by stomach-tube (gavage). Other methods, such as in capsules, in drinking-water, or mixed with diet, either introduce uncertainties into the records or, by altering appetite, may affect actual intakes (5, 69, 165, 169, 173). Then,

too, animals are commonly fasted before the dose (24); this itself may be embryotoxic or may alter a teratogenic effect of a test substance.

(2) Air pollutants and anesthetics have been tested by inhalation (47, 88). Direct "environmental exposure" of rats on the property of a synthetic rubber factory has been reported (97). But amounts inhaled have been hard to measure reliably. When air concentrations of a test substance are increased, the animals may respond by breathing less deeply or less often. So, reliable dose-response curves are hard to obtain. Intranasal instillation does not avoid this difficulty.

(3) Skin applications must be protected from molestation by the test animals if results are to be valid (165). Molestation transfers some of the dose to the oral route. But a small test animal has many times the specific skin surface area of a woman. Clipped or shaved skin may respond differently than bare human skin, and the animal's thermoregulation may be disturbed. The smaller the prepared area, the more concentrated the dose must be. These considerations limit the interpretability of test doses given to animals by this route.

(4) Ocular instillation is used to test eye medications (145). Some air-borne environmental chemicals may be suitable for testing by this route, but lacrimation would introduce uncertainties.

(5) The parenteral routes are mainly subcutaneous, intramuscular, and intravenous. Intraperitoneal injection is sometimes used as an alternative to gavage, but microorganisms, pH, enzymes and other conditions of the gut lumen are bypassed. Intra-amniotic, intrauterine and yolk sac injections, and placing of drug treated millipore filter on the amnion or the placenta are also current research procedures (26, 128); these bypass the metabolic systems of the mother and the placenta, and could give misleading information if used to evaluate environmental hazards. Also, the procedures in themselves can be teratogenic (80).

It may be noted that none of these routes adequately reproduce the conditions of uptake by man of an environmental chemical, when a reliable teratogenic dose-response curve has to be obtained for test animals. The question raised by Kalter, whether it is logical to make the attempt, seems to be justified; this survey does not answer it. It may be that any route that produces equivalent post-hepatic plasma levels of the test substance or an appropriate metabolite in the female animal could satisfy criteria for validity of the test. But evidence is lacking, and meanwhile, the principle of approximation is accepted.

c. Duration of treatment. Human exposure to environmental chemicals varies from continuous at low levels to intermittent at high levels.

Continuous exposure of a human female will usually have started long before her pregnancy. But, for example, a newly married female may move into an exposure area and become pregnant immediately. Continuous exposure may produce either cumulative effects or, if metabolic tolerance develops, reduced effects.

Cumulative effects tend to prolong exposure of the conceptus beyond the period of maternal exposure. Metabolic tolerance, on the other hand, tends to reduce the period of exposure of the conceptus. An investigator may be able to predict which sort of effect is more likely from preliminary test data on metabolism of the substance, especially in pregnant females.

Intermittent high-level exposure of a pregnant female may, depending on how she metabolizes the substance, deliver a short, sharp challenge to the conceptus, or a more prolonged challenge at a lower level if, for instance, the substance is rapidly stored in body fat.

If human exposure is expected to be low-level, continuous, and everywhere, without episodes of higher exposure, a teratogenicity screening test may not need to include short-term exposures of test animals. On the other hand, at least one group of test animals should probably always be screened for cumulative effects by continuously administering the substance throughout pregnancy.

In general, however, the uncertainties suggest that durations of treatment with a substance should rarely be based on expected human exposures to the environment, but rather should be directed to revealing thresholds of sensitivity at its greatest in the test animals.

Most protocols and guidelines for teratogenicity tests (69, 163, 165), particularly those concerned with regulation, specify that the substance shall be administered to the test female throughout the period of major organogenesis (Tables 10, 11 and 12). This time period, the "critical period of organogenesis" differs among the species; it is in part dependent on the duration of gestation and is the period during which the embryo is highly susceptible to teratogenic insult. During the predifferentiation period in early gestation and following differentiation, the conceptus is generally resistant to production of congenital malformations although embryonic death and/or abortion may occur. Sometimes teratogenic treatment reveals more than one period of maximum effect (Table 12).

Administration of drugs that dissolve and are absorbed, metabolized and excreted rapidly can be confined to the "critical periods." But drugs that are absorbed or metabolized more slowly will have effects that extend beyond this period, and result in apparent absence of teratogenic effect if tested only during the critical periods. This factor represents one of the most common pitfalls in teratogenic testing and accounts for the observation of teratogenic activity outside the critical period. Drugs acting in this manner (actinomycin D, cyclophosphamide, striptonigrin) generally require activation in situ for a period of time before they can be effective (18).

Wilson (163, 165) points out that continuous treatment during organogenesis may produce maternal adaptations that decrease the effect on the conceptus (Tables 13 and 14). He suggests the use of several treatment spans: one group to be treated continuously to reveal any cumulative effects, other groups to be treated for 3-4 days at a time to avoid adaptations of maternal enzyme systems (Table 15). Table 16 shows an extension of this idea that is comprehensive and beyond the scope of this report.

To avoid attribution of any pre-implantation defects to the test substance, the start of pregnancy is observed and recorded. For example, in rats and mice a vaginal plug or sperm in the vaginal smear can be seen after mating. The morning this evidence is first seen marks the first day of pregnancy. The sensitive period of major organogenesis is then worked out, using tables.

Other periods of sensitivity to teratogenesis have received less attention in the past, but may prove more important in the future. One of these is late-stage teratology of the CNS (69). Recently Langman and his associates have been developing information on the timing and morphology of sensitivity in specific areas of the CS (87, 123, 136). However, protocols for the routine screening of environmental chemicals for late-stage CNS teratology have not yet been fully developed.

d. Dose vehicle. The vehicle is more important than is often supposed. Arguments that its properties do not matter so long as control and treatment groups all receive like amounts have been shown to be untenable (5).

The Canadian guidelines for teratogenicity testing (5) discuss the vehicle at length. The major requirements can be summarized as follows:

- No alteration of the chemical properties of the test compound.
- No serious influence on absorption, distribution, metabolism or retention of the test compound.
- No increase, decrease, or other alteration of toxic properties of the test compound.
- No local or systemic properties of its own.
- No effect on consumption or utilization of foods or drinking-water.
- No interactions with other conditions of the proposed test.
- If possible, the vehicle should be comparable with that to which humans are expected to be exposed in practice.

  If possible, the vehicle should dissolve the test substance.

Table 10

Episodes in the Reproductive Cycle of Mammals (163)

1. Preservation of the germinal line
2. Gametogenesis
3. Release and transport of gametes
4. Fertilization
5. Cleavage and blastocyst stages
6. Implantation
7. Metabolic changes in pregnant dam
8. Embryonic period - organogenesis
9. Fetal period - histogenesis, growth and beginning functional maturation
10. Placenta, maternal-conceptal relations
11. Birth and postnatal adjustments
12. Lactation and maternal care
13. Postnatal growth, functional maturation

Source: Reprinted from "Critique of Current Methods for Teratogenicity Testing and Suggestions for Their Improvement" by J.G. Wilson, in Methods for Detection of Environmental Agents That Produce Congenital Defects, T.H. Shepard, J.R. Miller and M. Marois, Eds., by permission of Excerpta Medica - North Holland Publishing Company. Year of first publication, 1975.

Table 11

Gestation Times for Some Laboratory Animals and Man (128)

Species	Mean duration of gestation days
Chick	21
Rat	21
Hamster (golden)	16
Mouse	19
Rabbit	31
Ferret	43
Cat	63
Dog	63
Guinea pig	68
Pig	114
Sheep	150
Monkey (rhesus)	168
Monkey (baboon)	175
Armadillo	225
Human	278

Table 12

Critical Periods of Organogenesis in Animals (128)

Species	Critical period[a] days
Chick	1-3[b]
Rat	9-17
Hamster (golden)	4-14
Mouse	7-16
Rabbit	8-21
Ferret	8-28
Cat	5-58; 5-15 most favorable
Dog	1-48; 8-20 estimated
Guinea pig	11-20
Pig	12-34
Sheep	14-36
Monkey (rhesus)	20-65; 22-30 most susceptible
Monkey (baboon)	22-47
Armadillo	1-30
Human	20-55

[a] Period of embryological organogenesis or period of known susceptibility to teratogens.

[b] Taken from reference 128.

Table 13

Some Chemical Agents Known to Influence Rates of Metabolic Degradation of Themselves and/or Other Repeated Dosage[a] (165)

Increase metabolic degradation	Decrease metabolic degradation
Barbiturates	SKF 525-A
Thyroxine	Chlorthione
Some insecticides (DDT, chlordane, aldrin, dieldrin, heptachlor)	Iproniazid
Some tranquilizers and antipsychotics (meprobamate, Librium, chlorpromazine)	Metopirone
Some antihistamines (chlorcyclizine, diphenhydramine)	Actinomycin D
Several hypoglycemic agents	Puromycin
3,4-Benzypyrene	CCl_4
3-Methylcholanthrene	Triparanol
Steroid hormones	Chloramphenicol
	Any that competitively inhibit catabolic enzymes

[a]From various sources.

Source: Reprinted from Environment and Birth Defects by J.G. Wilson, by permission of the Academic Press. Year of first publication, 1973.

Table 14

Ways in Which Repeated Treatment Prior to the Peak Susceptible Period of the Embryo May Produce Misleading Results (165)

Time of treatment	Primary effect	Secondary effect capable of altering test results
1. Before implantation	Interference with implantation	No issue
2. Early organogenesis	Early embryonic death	No issue
3. Before peak susceptibility	Induction of catabolizing enzymes	Reduced blood level during susceptible period
4. Before peak susceptibility	Inhibition of catabolizing enzymes	Increased blood level during susceptible period
5. Before peak susceptibility	Liver pathology or reduced function	Increased blood level during susceptible period
6. Before peak susceptibility	Kidney pathology or reduced function	Increased blood level during susceptible period
7. Before peak susceptibility	Saturation of protein-binding sites	Increased blood level during susceptible period

Source: Reprinted from Environment and Birth Defects by J.G. Wilson, by permission of Academic Press. Year of first publication, 1973.

Table 15

Numbers of Pregnant Rats Treated At
Consecutive 3-day Periods (165)

Dose level	Gestation Days		
	9-11	12-14	15-17
X^a	10	10	10
X/2	10	10	10
X/4	10	10	10

X^a - highest tolerated dose over 10-day period, or half of adult LD_{50}, or other appropriate effect-level.

Source: Reprinted from Environment and Birth Defects by J.G. Wilson, by permission of the Academic Press. Year of first publication, 1973.

Table 16

Types of Test for Developmental Abnormality
Depending on Duration of Treatment (163)

1. Throughout reproductive cycle - whole generation(s).
2. Throughout pregnancy - conception to term.
3. Throughout organogenesis, primitive node to palatal closure.
4. Short-term (3-4 d) sequences during organogenesis.
5. Aimed at specific events - mutagenesis, postnatal, etc.
6. Combinations of above, e.g., 3, 4, and 5.

Source: Reprinted from "Critique of Current Methods for Teatogenicity Testing
and Suggestions for Their Improvement" by J.G. Wilson, in Methods for Detection of Environmental Agents That Produce Congenital Defects, T.H. Shepard, J.R. Miller and M. Marois, eds., by permission of Excerpta Medica - North Holland Publishing Company. Year of first publication, 1975.

Instances are cited (5) of seemingly unimportant properties of vehicles that made test data "difficult if not impossible" to interpret. Dimethyl sulfoxide enhanced penetration of some test substances through the skin and altered their distributions in the body. Even the provision of distilled drinking-water instead of deionized or tap water affected litter size, weight gain, and water-consumption of pregnant rats.

Information may be lacking on the extent to which a vehicle satisfied these requirements in the quantities to be used during a teratogenicity test. In cases of doubt, a vehicle should be evaluated separately against a group of completely untreated controls under the conditions of the proposed test (5).

Others have noted that a vehicle should maintain the physical state, concentration, and other requirements of exposure to the test compound, and if water is used, it should have no contaminants such as chlorine or phenols (71, 167).

e. Controls. The purpose of all controls in teratogenicity testing is to help the investigator answer the question, Were these malformations that I have observed caused by the test substance, or not? Often this question is difficult or impossible to answer, however good the controls.

One reason for the difficulty is the range of normal variants and "spontaneous" defects found in man and most test animals, from which induced defects need to be distinguished. As Wilson points out (165), "clearcut embryotoxicity... is readily recognized but clear-cut cases are by no means the rule" because most highly embryotoxic agents will have been eliminated before the stage of testing for teratogenicity.

A more important reason for good controls, according to Kalter, is the possibility that unknown variables in the test animals or their environment may modify the teratogenicity of the test substance, making it more or less teratogenic, or making the animals more or less sensitive, or both.

Therefore, the most useful control is the investigator's background knowledge of the strain of animal employed - the "historical" control group that comprises the accumulated data on all defects observed in the past in untreated animals of that particular strain (69, 71, 167). If, for example, anophthalmia had been observed at the rate of 1 in 200 offspring, and no cases of spina bifida had been seen, in the historical controls, and then if one case of each were to be found in a treated group, those two cases would be interpreted differently (165).

The quantity of the historical control group is an important factor in selecting the species and strain of test animal. Usually, a species or strain with adequate historical controls will be preferred over one that seems developmentally more suitable but lacks historical controls. This, for example, is one reason why mice rather than guinea pigs have been used for late-stage CNS teratogenicity tests.

Equally necessary are concurrent "negative" controls, groups of animals that are similar to the test groups and are treated similarly but without the test substance. Concurrent negative controls have three functions:

- They are compared to groups treated with the test substance.
- They are compared to the historical control in order to detect the presence of unexpected variables.
- They contribute to the accumulated historical control for future tests.

All concurrent control groups are matched meticulously with test groups for:

- Number of animals in group.
- Characteristics of the animals (age, size or weight, history, health)
- Test environment (all aspects of husbandry, amount and times of handling, and identity of handler).

The purpose is to reduce the known variables to one: the differences that need to be detected and measured.

The treatment of a test group will in part determine that of the concurrent negative control group, and may itself be partly determined by control requirements.

An important instance is attention to the test vehicle (5, 71, 165, 173). If a concurrent negative control group is to be compared with a group receiving the test substance, the controls will receive the vehicle comparably with the treated group as to amounts, times, route(s), and all other particulars. If omission of the test substance alters the quantity or character (e.g., volume or caloric density) of what is administered, the change is not rectified by automatically increasing the amount of vehicle, but alternatives are considered (e.g., a known inert placebo).

If activity of the vehicle is in question, the same principle applies, and an untreated control group is compared to the vehicle treated group which now acts as a test group. The concurrently untreated group is then compared with the historical control. In this format there are essentially three tiers of comparisons:

- Substance-plus-vehicle versus vehicle alone.
- Vehicle versus no treatment.
- No treatment versus historical control.

This set of comparisons can be expanded or condensed according to need, so long as the need is understood to be for maximum validity.

As for numbers of animals per group, the practice is to use 20 mice or rats, or 10 rabbits, at each dose level when the test substance is given throughout organogenesis. These numbers are halved for short treatments given at intervals during organogenesis or gestation (see Table 14). Recommendations for numbers of more expensive species vary (165).

In the past, groups of animals designated as concurrent positive controls used to be given doses of some compound that for them was a known teratogen. This was considered to indicate how sensitive the particular strain of animals was to teratogenic influences in general. The degree of sensitivity to a known teratogen is no sure guide to the degree of sensitivity to a different test substance.

3. The Test Environment

Many aspects of the test environment have been discovered to affect the development of the conceptus and its sensitivity to teratogens. They will be considered under the headings (112) of (a) diet, (b) caging, (c) climate, (d) stresses. When test conditions vary among laboratories, or from time t time in a laboratory, test results become hard to interpret. At present, more constancy and uniformity is needed. This implies high standards of animal care and laboratory records (5).

a. Dietary factors that can cause defects or modify effects of teratogens include:

- o Inappropriate feeding times, e.g., fasting (5, 69, 165).
- o Composition of drinking-water (71, 167).
- o Balance of the major nutrients (5, 118, 165).
- o Composition of chows can vary among sources and among batches from any source, affecting palatability and intake; thus composition is often a trade secret. Knowledge is needed, in order to maintain constancy.
- o Decomposition of diets, especially chows, with age or spoilage.
- o Contamination with endogenous alkaloids (69, 77), with incorporated antibiotics (18, 71), with pesticides, dirt, and excreta, or with microorganisms.
- o Serious deficiencies or excesses of vitamins (5, 46, 69, 159, 165).
- o Mineral deficiencies, e.g., Zn, Mg, Mn, I, Cu, (16, 63, 71, 159, 165) or excesses, e.g., I, Hg, Pb (1, 16).

Reviews are available of nutrients known to affect development of the conceptus (62), of methods of evaluating animal feeds (4, 30), and of the nutritional requirements of certain test animals (3, 46). It is very important to control the quality of the diet. For example a 30% reduction in food intake in mice produces effects on the teratological outcome of the results, while a 50% reduction is required in the rat to produce similar results (9).

b. Caging is mainly a matter of quality (5, 173); a poor surface can harbor microorganisms or yield variable amounts of unidentified minerals. Bedding is sometimes eaten or chewed, and if treated, it may contain substances that can affect test results, e.g., ethylene oxide, organochlorine compounds.

c. Climate includes many potentially teratogenic factors (112):

- Extremes of temperature, whether of heat or cold, can produce malformations or potentiate other teratogenic influences (5, 25, 32, 69, 107, 112, 165). Drafts are to be avoided; the temperature should remain even throughout an animal room.
- Relative humidity should be constant or recorded, extremes or sudden change can stress test animals or affect the behavior of pathogens.
- Extremes of high or low atmospheric pressure or oxygen tension can be teratogenic (5, 69).
- Light should be sufficient, of appropriate spectra quality, and evenly distributed. Light cycles must be regular for reproducible results (165), and they can affect determinations of gestational age (69). If cycles cannot be controlled, there may be seasonal differences of reproduction (173).
- Some test animal species have shown seasonal sensitivities to teratogens (69, 165).
- Any possible source of harmful radiations should, of course, be eliminated (69, 165).

Sudden or excessive noise should be avoided or recorded (5, 117).

Environmental odors can affect food intake and health of test animals.

d. Stresses of many kinds can cause or enhance teratogenicity (112):

- Overcrowding (69, 173), or immobilization (14), or excessive solitude.
- Odor of a strange animal (15).
- Physical trauma (165).
- Excessive, clumsy, or unexpected handling by laboratory staff (5, 78).
- Pathogenic microorganisms and parasites (5, 14).
- Medications (69, 165), and chemicals applied to the animal rooms (165).
- Psychological stresses (112) which often are hard to define but include failures to follow routine.

The above list is not exhaustive. Usually all possible sources of stress cannot be totally eliminated , but they can be recognized, recorded, and - most important - unequal incidence among the animal groups can be avoided.

4. Observations

The four signs of embryopathy are death, growth retardation, malformation, and functional deficit. Malformations may reflect cytotoxicity or brief, local episodes of retarded growth; the observed categories are external, skeletal, and visceral malformations.

To look for all of these signs, procedures must be timed carefully, so timing will be considered before methods for observation. But the first question, because it affects the approach to observation, is whether any condition of pregnant females can be teratogenic of itself.

a. Maternal effects. Can a conceptus be malformed indirectly, as a result of toxicity to a pregnant female, when no toxic substance or metabolite reaches the conceptus? The answer seems to be qualifiedly negative (165). According to Wilson, no examples of this have been seen in test animals, although some observations in man have been interpreted as secondary malformations. Secondary fetal deaths have been reported, for example, after serotonin was given to pregnant rats, but there were no malformations nor potentiation of teratogenicity of other substances (51).

If a pregnant female receives a diet deficient in certain respects, this can produce a malformed conceptus, but such "teratogens" are not toxic substances or metabolites. Nutritional antagonists have been hypothesized but, according to Wilson, have not been demonstrated. Malformations attributable to position effects in multigravid uteri are considered marginal or trivial.

So, according to Kalter and Wilson, secondary embryopathy results either in death of the conceptus or in retarded growth observed at birth and reversed, usually, by proper feeding of the neonate. Teratological effects would not be reversible. When a pregnant female fails to gain weight at the expected rate, secondary embryopathy can be expected.

Another important factor is the age and parity of the animal. In polytocous animals such as rats, rabbits, mice and hamsters the reproductive capacity has been shown to decline with increasing age. In mice the highest rate of spontaneous malformations occurs in the first litter and then decreases rather significantly by the fourth litter. In rats and rabbits, on the other hand, the size of the litter decreases with age, but the incidence of spontaneous malformations is not affected by age and parity.

b. Timing. Pregnant females are watched throughout gestation for weight gain and for signs of ill health.

As a rule, the females are killed a few hours before expected delivery to avoid cannibalism of stillborn, moribund, or malformed offspring, and to count fetal deaths and resorptions (5, 14, 35, 78, 153, 165).

When postnatal observations on neonates or older offspring are needed, some females are allowed to litter normally (5, 35, 78, 153, 165). Cross-fostering may be needed if the treatment has interfered, for example, with lactation.

One exception has been reported to the general rule of looking for birth defects at or soon after birth. Neonatal CNS damage from absorption of free acidic amino acids is phagocytized rapidly, within about 24 hours (12, 89, 114), and animals are killed 2-5 hours after treatment (113). Such damage can occur transplacentally in late pregnancy (Olney, personal communication 1974), and examinations can be timed accordingly if this is suspected.

c. Preparations. The pregnant female is killed by any means that does not injure the uterus or its contents. Unusual methods, e.g., perfusion-fixation, are employed when specially needed.

The abdomen is opened; the uterus is inspected in situ for total implantation sites including all states of resorption. Corpora lutea are counted and compared with the total of implantation sites. Faulty implantation tends to scar if it survives for 24 hours or more. By these or other suitable methods (depending on the expected accuracy for the species), preimplantation losses are estimated.

The uterus is opened; living fetuses are distinguished from intact dead ones by color, movement, etc., and macerated ones are counted as resportions. In the living fetuses, the usual vital signs are assessed, and the umbilical cords are clamped or cauterized. After preliminary examination, these fetuses are blotted, weighed, sexed, and prepared for detailed examination.

Fixation in Bouin's fluid is recommended by Kalter for animals to be sectioned either by Wilson's razor-blade method or by microtome. Sectioned tissues are stained appropriately. The head is often detached from the body and sectioned coronally, while the body is sectioned transversely. If there is reason to suspect selective teratology of a particular system, e.g., nervous system, special preparations are made at this stage for appropriate histology.

d. External malformations. The preliminary examination is by eye and, if necessary, low-power microscopy as described in detail by Wilson (161). Information on what to look for in various test species has been given by

Warkany (158), Kalter (69), and Wilson (164, 165, 169), and in man by Potts (119). Gross structural features are inspected, and any deviations of proportion are measured and recorded. This examination is conducted by a scientist with in-depth, relevant knowledge of the species employed.

e. Skeletal malformations. One-third to one-half of the offspring are fixed in 95% alcohol for skeletal visualization by a technique using the alizarin red S stain (14, 23, 24, 38, 74, 129, 165). The method of Schnell and Newberne is preferred (129, 165, and see Figure 2). But alizarin stains only ossified bone, and rodent skeletons will still be largely cartilaginous. Methylene blue may be used to stain rodent cartilage, and the Noback modification of Van Wijhe's method is recommended (111). Large animals have been examined by radiography (165, and see Figure 3) but good visualization may pose problems. The examiner should distinguish major or harmful malformations from normal skeletal variations (71, 167).

f. Soft-tissue malformations. Large animals, e.g., the 100-day monkey fetus, are examined by standard autopsy procedures. Any local or general growth retardation is revealed when organ measurements are compared with normal values (see Table 17) (165). Small animal offspring not used for skeletal examination are fixed in Bouin's fluid or other nonhardening fixative that decalcifies the skeleton. They are dissected by standard methods to reveal soft-tissue abnormalities. Wilson's freehand razor-blade technique (169) is more often used, both for regulatory tests and for research (24, 35, 38, 78, 151, 173). Some examiners section by microtome, or dissect the major viscera individually after fixation. A portion of liver is usually examined separately. Slices of tissues are compared with reference sections or with, e.g., Wilson's Atlas of Sections.

If there is reason to suspect that a substance may selectively damage a particular system, such as the nervous system, the examiner chooses an appropriate technique that will reveal histological changes - for example, serial sectioning. If special fixation is needed, or special times of observation, for example a need to kill test animals by perfusion-fixation, a sufficient additional group of animals is reserved for this.

Table 17

Weights and Measures on 15 Control Monkey Fetuses Removed by Hysterotomy on Gestation Days 99-101 From Untreated Rhesus Females (165).

Fetus	Sex	Body Weight (gm)	Amniotic fluid (ml)	Weights of major viscera (gm)						
				Brain	Eyes	Lungs	Heart	Liver	Kidneys	Spleen
13d	F	133	70	20.6	1.10	2.4	0.80	6.3	1.29	0.18
16e	M	154	135	22.5	1.02	2.6	1.00	6.1	0.86	0.22
29d	M	149	150	23.4	0.96	2.6	1.10	5.2	1.25	0.33
41c	M	151	125	-	0.96	3.5	1.10	5.4	1.12	0.40
40e	M	148	130	21.1	1.16	2.8	1.00	5.3	1.06	0.34
47c	F	119	95	15.3	0.92	3.0	0.70	4.5	0.78	0.26
48c	F	142	84	25.4	0.96	2.7	1.08	5.8	0.90	0.24
60c	F	130	105	21.0	0.98	3.1	0.90	4.5	0.97	0.30
75b	M	125	85	18.0	0.90	2.5	1.33	4.8	0.91	0.25
39d[a]	M	143	85	23.1	1.06	2.7	1.10	-	0.96	0.20
70b	M	159	71	22.2	1.23	3.2	1.30	7.0	0.98	0.35
94a	M	146	93	21.0	1.08	3.0	1.26	6.1	0.97	0.35
98a	F	155	86	21.1	0.93	3.5	1.50	6.3	0.91	0.33
77c	F	154	84		1.11	2.9	1.09	5.5	0.80	0.33
58f	M	143	47		1.05	2.2	1.08	5.1	0.82	0.29
Mean		143	96		1.03	2.8	1.09	5.2	0.97	0.29
Standard deviation		12	28		0.10	0.4	0.20	0.8	0.15	0.06

[a] The mothers of this and the following 5 fetuses were given a vechicle, 15 ml of 0.3% tragacanth daily by stomach tube, days 20-45 of gestations.

Source: Reprinted from Environment and Birth Defects by J. G. Wilson, by permission of the Academic Press. Year of first publication, 1973.

Figure 1

Cleared, alizarin-stained monkey fetus removed by hysterotomy from an untreated rhesus female on day 100 of pregnancy. These specimens allow excellent visualization of the ossified skeleton but require 2 to 3 months for preparation. The sternum was displaced forward during postmortem examination of the thoracic viscera (165).

Source: Reproduced from Environment and Birth Defects by J.G. Wilson, by permission of the Academic Press. Year of first publication, 1973.

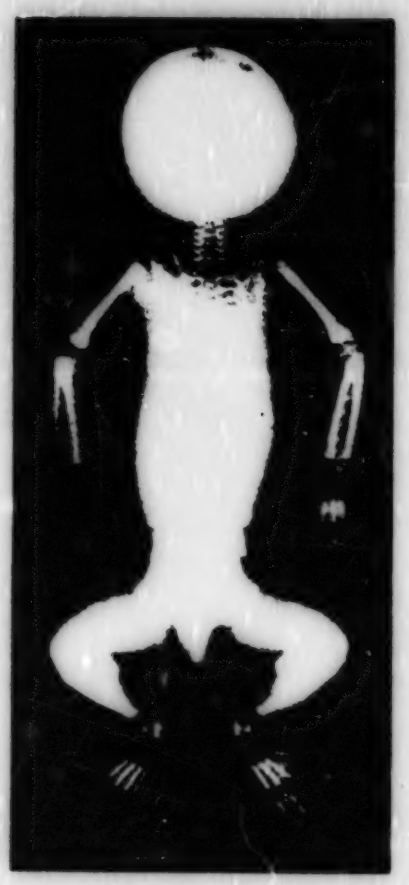


Figure 2

Radiograph of a monkey fetus removed by hysterotomy from an untreated rhesus female on day 100 of pregnancy. Although the definition of skeletal structures is not as clear as that in Figure 2, X-rays of this type have proved adequate for evaluation of the skeleton, and they have the advantage of being made quickly and inexpensively (165).

Source: Reproduced from Environment and Birth Defects by J.G. Wilson, by permission of the Academic Press. Year of first publication, 1973.

There is increasing, legitimate concern about possible postnatal functional defects not revealed by these procedures or by test animals that are born in a less developed state than the human. These defects are beyond the scope of this survey but are referred to briefly in the concluding sections.

g. Microscope observations. Freehand razor-blade sections are usually examined by dissecting microscope (29, 164). Microtome sections of whole animals or selected tissues are examined by ordinary light microscope (14, 24, 153, 157, 158, 173). Localized lesions and ultrastructures are examined by transmisson or scanning electron microscope (7, 131).

h. Biochemical observations. Some blood dyscrasias, congenital errors of metabolism, specific membrane defects, and other molecular abnormalities (171) may under some circumstances arise by teratological processes. These methods are no doubt useful for the study of teratogenesis; however, the extent to which they are suitable for screening purposes is an open question. However, as yet there is no sub-discipline of molecular teratology, and testing methods are beyond the scope of this survey.

i. Other observations. Subtle damage to reproductive systems may be detected by multi-generation tests, but in 1970 the FDA Advisory Committee warned of hazards in interpreting the results (42). Such tests are beyond the scope of this survey.

5. Interpretation of Results

People vary in their development, and so do test animals. Variability is easier to interpret if it is measured. Most laboratories draw constantly from the same few colonies of test animals. Thus they can accumulate records of anatomical variations, rates of intrauterine death, malformations, and growth-retarded offspring that occur in untreated control groups. These records tell them more than can any one concurrent control group. Some problems of interpretation are discussed below.

a. Growth retardation. Retarded growth in the uterus may be inferred if a fetus or neonate weighs more than two standard deviations below the mean of the matched controls. Unexplained cases occur in man (93) and undoubtedly in other mammals too. Localized growth retardation, e.g., of the palatal shelves that results in cleft palate, is a different phenomenon from the overall growth retardation that is the subject of this subsection. However, the two phenomena can overlap or coexist. Some chemicals can produce both independently, or one can lead to the other, or either can result in intrauterine death.

Overall growth retardation usually involves delayed development of the ossified parts of the skeleton. It is generally seen at doses that are toxic to the maternal organism, to the fetus or both. Stunted fetuses are often malformed and are entirely different from "runts" which are caused by placental fusion and position of the fetus in the uterus, at least in mice.

To interpret a finding of growth retardation, two questions are often asked: Is the effect permanent or reversible? Is it undesirable? Growth can be permanently impaired by, for instance, prolonged exposure to low levels of radiations (66), which are clearly undesirable. On the other hand, reversible growth retardation can result indirectly from maternal toxicity or underfeeding that interferes with the supply of some or all nutrients to the fetus.

The point at which relative 'smallness' becomes undesirable is not well defined. On the one hand, it has been known for more than 40 years that a less-than-maximal growth rate and adult size can conduce to health and longevity (e.g., 49). On the other hand, a recent 12-generation study of rats fed a marginally protein-deficient diet produced overall growth retardation with unevenly altered proportions of various viscera to body weight and with behavioral changes; not all of these deviations could be reversed by transfer of animals to a full diet, and some of them were considered undesirable. The authors concluded that this study was extrapolable to man in principle, if not in detail (140).

Because these questions generally cannot be answered within the time available for routine teratogenicity tests, an animal weighing at least two standard deviations below the mean of matched controls is usually assumed to be undesirably growth-retarded.

b. Malformations. In human teratology a malformation can be defined as an anatomical abnormality that either kills or requires surgical intervention to maintain a normal existence, or is a cosmetic defect that reduces the quality of life. When cadavers are dissected, circulatory system variants are found in a continuum from near-infinity in the microcirculation to large, infrequent deviations of the aorta. Most of these caused little or no discomfort or restriction of function during life. The same applies to supernumerary spleens. The concept of detriment to the individual raises questions that are unanswered except according to particular circumstances. For example, detriment has been defined as: "actual or potential cause of disease, disability or psychologic impairment"; alternatively, as "actual or potential hazard to survival, health, or usual activity."

Thus, detriments to people often do not correspond with detriments to test animals. For example, cleft lip is a cosmetic detriment to man but a lethal detriment to a rodent which thereby cannot suckle. On the other hand, polydactyly and ectrodactyly are classed as malformations in test animals because they are malformations in people, while agenesis of tail is normal in people but a detriment to some species of test animal. Skeletal variations of ribs and thoracic and lumbar vertebrae are not considered abnormal in mice, rats, rabbits, hamsters, guinea pigs, or rhesus monkeys even if frequencies increase after treatment with a known teratogen; control treatments (inert injections) also sometimes increase the frequency. Wilson suggests that such increases may indicate that embryotoxicity is "being approached" or that irregular ossification "may simply reflect growth retardation," so that body weight records are needed for interpretation (165).

Therefore it is hard to define what is a malformation in a test animal, especially in terms of what might be a malformation if it were found in people. Defensible criteria, e.g., limits of magnitude or of statistical rarity, do not seem to have emerged.

One suggested criterion has been that a treatment shoudl increase the frequency of known 'background' (accumulated control) variations (2, 64, 86, 125, 126). However, some teratogens may not do this. When a variant or malformation has been reported in accumulated controls, a significant increase of the same in treated animals is cause for suspicion of teratogenicity, but no more.

An unequivocal criterion of ceratogenicity, regardless of control findings, is a dose-related increase of malformations in treated animals.

c. Death in utero. The proportion of dead implants per female appears to be a more appropriate measure of fetal death than the number of dead implants per female because the analyses based on proportion take into consideration the total number of implants per female (57).

The relationship between deaths and malformations can be complex and according to Wilson (165), is not always interpreted correctly. This is particularly evident in cases where both manifestations tend to appear at comparable dose levels during organogenesis and have parallel dose-response until, at high doses, deaths replace malformations (e.g., cyclophosphamide) (18).

However, there are some drugs which produce embryolethality but not malformations. This suggests that embryolethality and teratogenesis are two separate phenomena produced by different modes of toxic action (e.g., methotrexate, 6-aminonicotinamide) (18). Obviously, selection of the proper dose is of paramount importance. If the dose is too high, a 100% lethality will occur. If the dose is too low, no effect will be produced on the fetus and the results will provide a false sense of security about the drug's safety.

Any increase of either deaths or malformations denotes toxicity. When rhesus monkeys are given toxic doses, deaths always seem to occur more readily than malformations. If this also applies to people, it "would in most circumstances be regarded as preferable" to an increase of malformations (165).

d. Positive controls. These are controls treated with a known teratogen for comparison with the substance whose teratogenicity is unknown. They ought not to be needed, for three reasons:

- The investigator should already know the general sensitivity of his animals,
- Sensitivity to one agent is no guide to sensitivity to another; for example, rats are sensitive to trypan blue and to late antagonist but not to thalidomide or cortisone, and
- The range of sensitivities differs among strains of test animals, including man.

Rather, the aim should be to find embryotoxicity in two or more unrelated animal species and, if animals are available for a positive control group, it would be more profitable to add them to a treatment group (165).

e. The "litter effect". Often malformations are concentrated in some litters, sparing others (65, 170). This usually denotes maternal toxicity (40) but may also reflect the genetic similarity of littermates. The question of how to interpret such clustering is much debated, but it is complex and there are not enough facts to provide answers.

Many statisticians prefer to count the litter (pregnant female) rather than the fetus as the basic experimental unit because they feel that the per litter evaluation, which is based on the average reaction within the litter, automatically includes the reaction of the individual fetus (5, 138, 162). The number per litter tests has been recommended by Haseman and Hogen (57) for analysis of teratologic incidence data.

Other investigators prefer to count individual fetuses, pooling the total in a treatment group or comparing malformation rates per litter among groups (10, 60, 165). This approach is misleading and may seriously exaggerate the significance level (57).

Besides the selection of the litter as the experimental unit, there are other factors involved in the appropriate statistical analysis of the

teratologic data. These include the number and size of the test population, distribution of the variables of interest, nature of comparison desired, the magnitude and degree of litter effect, and fetal age and size at the time of treatment (85, 90).

Questions have been raised about how much difference is involved in the use of different methods, but they have not as yet been resolved (10, 138).

f. Sources of error. If results are not clear-cut, the first impulse is to repeat the test. This will be of little use if the first test contained errors that are not corrected. Such errors may be hard to recognize because of unknown variables, the difficulty of controlling known variables, and disappointing reproducibility (17). Possible causes of poor reproducibility that require vigilance, from the design of a test to its interpretation, include seasonal fluctuations of effects, spontaneous mutations, spontaneous or infectious abortions, inaccurate timing of fertilization, and irregular administration of test doses (165).

g. Statistical methods. For interpretation of teratologic results, the following statistical methods are employed: two-way analysis of variance, regression analysis and tests for best fit. Analyses based on individual fetuses employ the chi square test, Fisher's exact test and the t-test. It is believed that more rigorous statistical analysis is required to maximally evaluate the teratologic results and risks.

B. In Vitro Systems

In vitro systems, including tissue cultures and avian eggs, have contributed to basic research on cellular mechanisms of teratology, but have virtually no role as short-cuts in screening chemicals for potential teratogenicity for man. Authorities do no agree on the value of avian eggs as coarse indicators of potential teratogenicity. In this test system untreated eggs can exhibit teratogenic effects from careless handling. In addition, embryos have no way of excreting the test compound or its metabolite. For

example, malathion was found to have teratogenic effects (monster formation) when evaluated by this test method, but there were no adverse effects when malathion was tested on mammals or in humans exposed during its manufacture (79).

However, cultures of embryos and cells may become useful in the future for studying metabolites of environmental chemicals that arise in mammals. Conceptually, the questions that in vitro systems could answer await development; this requires parallel correlative studies with particular chemicals in vivo and in vitro, and few such studies have been reported. Table 18 summarizes a selection of reported in vitro systems.

Table 18
Some In Vitro Systems That Have Been Employed in Teratology

System	References
Embryos, vertebrate	
Mammals:	
Rat	102, 106, 121, 154
Mouse	5, 82, 102, 105
Birds:	
Chick explanted embryos	56, 83, 103, 156
Amphibians:	
Frog	20, 21
Toad	12, 91, 132
Newt	19, 20
Salamander	133
Fish:	104
Embryos, invertebrate	
Echinoderms:	
Sea Urchin	52, 117
Sand Dollar	75
Insects:	
Fruit Fly	15
Organ Cultures	
General:	5, 14, 151
Mouse long bones	127
Mouse limb buds	109
Mouse tail vertebrae	175
Rat eyes	8
Chick embryo organs	14
Tissue Cultures	
Chick embryo tissues	14
Microorganisms	
Protozoa:	
Flagellates	14
Algae:	
Desmidiaceae	116

IV. CURRENTLY USED METHODOLOGIES

No one routine is considered reliable enough for use without question. In 1966, the FDA and in 1967, the WHO developed principles for testing chemicals for teratogenicity. The recommendations issued constitute only "guidelines" to methods (155). The onus is on the manufacturer, who has prior and superior knowledge of his product, to propose methods that the FDA will then approve or reject. The guidelines have been developed for new drug applications and have received comment on this basis (35,78). No comparable methodologies have been developed specifically for environmental chemicals, and to extrapolate from the FDA guidelines one should bear in mind that conditions of exposure and the reasons for it may be very different.

The FDA guidelines (155) divide testing into three "segments": I. Study of Fertility and General Reproductive Performance; II. Teratological Study; III. Perinatal and Postnatal Study.

Segment I covers the entire span of reproduction in both sexes and is considered a pilot series, to "serve as a guide for subsequent studies in depth" (155). Kelsey (78), an FDA scientist, noted that this segment is often run along with chronic toxicity tests for convenience, using the same test animals. She described a typical protocol: females are dosed from two weeks before mating until they are killed or their offspring are weaned, and males are dosed for 68-80 days before mating. Half of the females are killed on day 13 and examined for number, distribution, and condition of embryos, implantation sites, etc. The others litter naturally and are observed for duration of pregnancy, litter size, stillbirths, and malformations of newborn. According to results, reproduction may be tested in the offspring, or the same females may be studied through a second litter. The guidelines themselves (155) give no protocols, but Segment I clearly has a wider scope than teratology.

Segment II is intended to detect both teratology and embryotoxicity (155). The guidelines restrict doses to the period of organogenesis and also envisage second-litter and second-generation studies. They include some precise instructions. At least two species should be used (mouse, rat,

and rabbit are used most often). Males are untreated. Mice and rats are dosed on days 5-15, rabbits on days 6-18. Cesarean section is performed 1-2 days before expected delivery.

Observed are: number and placement of fetuses, living and dead, early and late resorptions, and corpora lutea. Each fetus is weighed and examined for external defects. Internal defects are always looked for: skeletal defects using alizarin stain, and visceral defects by dissection or the Wilson method. Rat fetuses are randomized into subgroups for these searches. Rabbit fetuses are examined for all types of defect, and half of them may be incubated for 24 hours with emphasis on observing the 6-hour survival rate.

Kelsey commented (78) on the need to choose a species "known or believed to metabolize the drug in a similar fashion as man", and to use two or more dose levels including the maximal tolerated dose.

Segment III specifies (155) that animals should be dosed for the last third of pregnancy and afterwards until weaning. The studies therefore are beyond the scope of this survey.

The guidelines also mention a number of specific points. Segment II studies can be made more specific by dividing the dosage period into sub-periods that add up to the prescribed period. Chick embryo assays should be used "for ancillary" data only". At least 20 female rats or mice (some laboratories use 50) or 10 rabbits are required per group, and control groups should be "relatively large". Each laboratory should report its background data, i.e., accumulated negative controls.

Kelsey (78) pointed out that the FDA accepts studies conducted outside the United States if the investigator and laboratory are known and approved by them. She emphasizes the need to know more about environmental long-term, low-dose exposures, for which improved methodologies are required.

In 1973 the Canadian Department of Health and Welfare issued the report of a working party in the form of a review with recommendations (5). In general, these are parallel to the FDA guidelines, but they cover more ground, and there are some differences.

The investigator is advised to choose two out of three basic treatment schedules: on certain days of gestation, throughout gestation, or before and during gestation. The period of organogenesis is given as days 6-15 in mice and rats, 6-18 in rabbits, 4-14 in hamsters, 6-20 in guinea pigs, 7-35 in pigs, and 9-40 in monkeys. Two species are chosen: the rabbit, and the mouse, rat or hamster. The significance of placentation differences is treated as uncertain, and so are the effects of maternal age and parity, regarded as more important. In vitro test systems are not recommended, but existing data should be reported.

Much emphasis is placed on husbandry, in particular on ensuring that test animals are given an inert environment free from, for example, disinfectant residues. It is thought important to copy the expected route(s) of human exposure to the test substance, so far as possible. The dose vehicle is expected to comply with rigorous standards of inertness and to allow the dose to be given in solution. If "drug" combination effects are anticipated, these should be studied.

Observations generally follow the FDA guidelines with added emphasis on delayed effects and a recommendation to observe some animals till they are weaned (42). Background information on human malformations is required, and problems of extrapolation of animal data to man are discussed, as are statistical problems of interpretation such as litter effects. The report concludes with three model protocols given as typical examples rather than as specific recommendations: (1) for mammalian species generally, (2) for rabbits, and (3) for rats.

In Britain three dose-levels are reportedly used, and Robson commented that the additional level permits logarithmic spacing and also better separation of embryolethality from teratogenicity (122).

In summary, no standard methodology is required for regulatory tests of teratogenicity, but some general principles are recommended fairly uniformly. These include use of the rabbit plus one or more of several small rodent species, emphasis on background knowledge of the species chosen, dosage with at least two levels of the test substance, administration by a route similar

to anticipated human exposure, large negative-control groups and meticulous husbandry with special attention to a teratologically inert environment. Protocols are expected to vary according to the compound to be tested, and the responsibility is on the manufacturer to propose in detail according to his superior knowledge of the product. Teratology is not the first aspect of toxicity to be tested when a chemical is screened by the manufacturer.

V. SURVEILLANCE

Surveillance is the approach to meaningful information on the incidence of human malformations. It is the only usual source of data on frequencies of particular malformations in large human populations, and it can reveal much that is unexpected. As an indicator for human teratologic risks, surveillance is considered equal or superior to animals tests.

Four sorts of information are provided, according to the FDA Advisory Committee on Protocols for Safety Evaluations (42): (1) accumulated exposure data; (2) tissue samples from human surgery or autopsies; (3) notification by doctors; (4) epidemiological surveys especially in occupational groups. A problem has been that such information is acquired very slowly (54). Indeed, surveillance information can be reliable, comprehensive, complete, or prompt, but seldom has it been all of these at one time. So a choice of surveillance methods can be determined by one's priorities.

Trends in surveillance methods were reviewed in 1974 by Flynt (36). He reported that six priorities had emerged at a recent WHO conference: (1) monitoring; (2) epidemiologic studies; (3) registries; (4) detection of new syndromes; (5) education of doctors and the public; (6) public relations. Each of these priorities had its own requirements. For example:

(1) Monitoring required prompt information (within 1-2 months of birth) from selected sources. A source could be unrepresentative so long as it remained constant over time.

(2) Epidemiologic studies required large, mixed populations and complete reports. Promptness was not required.

(3) Registries were used for genetic studies, planning of health services, and follow-up casework. Only the casework required prompt reports.

(4) New syndromes were detected through accurate and complete reports from selected sources with adequate facilities. Promptness was secondary.

(5) and (6) Education and public relations were ancillary but were essential to any acceptance of a surveillance program by doctors and the public.

The history of surveillance in the United States was undistinguished until recently. In 1940, for example, birth certificates in New York State were printed on the back with a form on which to report certain malformations as Yes/No. By 1963, a 95% response rate was noted, but the information had not been regularly analyzed, and "shortcomings and biases" were admitted (92). After the thalidomide tragedy more attention was given to surveillance. By 1974, four programs were in operation: at Atlanta, Georgia; Albany, New York; Olympia, Washington; and Jacksonville, Florida.

The Metropolitan Atlanta Congenital Defects Program, started in 1967, is run jointly by the Center for Disease Control (CDC), the State of Georgia, and Emory University. In 1971, this program covered a population of 1.3 million (77% white) and was served by 21 hospitals and 4 chromosome laboratories. Malformed infants were registered mainly through visits to hospital nurseries, delivery rooms, and record departments. Stillbirth and infant-death certificates were supplied by the Georgia Health Department. Case-reports were collected from doctors who consulted the program's professional staff. For surveillance purposes, a malformation was defined as "any structural, chromosomal or biochemical abnormality in an infant diagnosed before his first birthday", but comments were that diagnosis before 7 days tended to be more reliable than later diagnoses (37).

A monthly report analyzing data gathered is circulated to those who contribute data. If clustering is seen, families are interviewed about possible environmental exposures. Since 1971, data have been monitored by computer, and since 1974, a CDC computer automatically prints out statistically significant increases above 'background'. By 1974, 23 defects had shown such increases, but no environmental causes had been identified. In 1971, the program cost about $50,000 a year and made people aware of malformation problems (36,37).

In 1975 Dr. J.W. Flynt told us informally that CDC is now processing data on 65,000 births a year from various programs throughout the United States.

Expected incidences of some malformations will be calculated from 1970-1973 data, when assembled. The automatic printout of significant increases will be used by CDC to offer help to state health departments in three different ways: (1) reviewing individual hospital case records; (2) interviewing families; (3) offering services of trained epidemiologists. In Dr. Flynt's opinion, there were still too many unknowns to permit the valid design of systems for automated data-analysis or the monitoring of special occupation groups of women, and such activity might be interpreted as invasion of privacy. However, highly organized societies outside the United States, such as Japan and Poland, have in the past welcomed American support for studies in special groups of their populations.

In Britain voluntary notification started in 1964, a midwife or doctor observing effects "in any way he wishes" and reporting to the General Register Office on a form listing 100 categories of birth defects (59). Information readily acquired included: locations of birth and of mother's residence, place of birth (home/hospital/other), birthdate, sex, living/stillborn, mother's age, and mother's total of live and stillbirths. Less often acquired were mother's name, history of multiple births, and infant's birth weight.

Currently, the UK Office of Population Censuses and Surveys receives notifications and alerts local authorities whenever the frequency of a defect rises by a significant 1% above the previous average. About one-half of the local authorities have followed up such notifications (160). Comments have been (59, 160, 161) that: efficiency of notification is less than complete, especially when the defect was lethal postnatally; a rise of defects is hard to distinguish from improved efficiency of notification; statistical approaches need more study; and therefore, such data must be interpreted with caution.

Informally we understand that British records have started to include the occupations of pregnant women with a view to future monitoring for teratogens. Surveillance of exposures to environmental chemicals does not seem to have been considered practical and an informal search for association between diseased potatoes and anencephaly (a recent topic in Britain) was negative.

In 1971 Miller at the National Cancer Institute, NIH, suggested (95) that a death-certificate registry for malformations, like that for cancers, might p[illegible]e useful. The data were routinely available but seldom coded. The quality of diagnoses would vary, but defects lethal in the first month of life (or during infancy) could be used. Fetal death certificates would be coded separately. These sources of information were, in the author's opinion, better than birth certificates. Other sources advocated by the author included Health Maintenance Organizations and veterinary observations; the latter had been useful in identification of mercury as the contaminant at Minnamata, and more recently in a Kentucky epidemic of lethal skeletal defects in pigs grazing on vegetables that had been sprayed.

The Federal Meat and Poultry Program of the Animal and Plant Health Inspection Service of the USDA does, in fact, operate a National Monitoring Program in order to ensure "a meat and poultry supply that is wholesome and contains no violative levels of any drug or chemical residue" (100). Sample animals for inspection are selected by computer, and tissues analyzed include fat, liver, kidney, and muscle. In 1973 over 110,000 analyses were performed on 18,000 animals for 47 different chemicals, and over 800 special surveillance programs were conducted after violations had been identified. The chemicals included teratogenic pesticides, and continual re-evaluation of the Program permitted rapid responses to new information.

In summary, the techniques under development appear to offer improvement in the surveillance of environmental teratogens. However, much better information can be obtained about total exposure from all sources if laboratory measurements to determine concentrations are combined with surveys to assess total intake. Amount of food, water, medication consumed, duration of work in different jobs or of residence in communities that differ in their potential for exposure to environmental teratogens are some of the factors that must be taken into account.

VI. ECONOMICS

The costs of teratogenic studies, as now required, are enormous. One estimate puts the minimum cost of a complete study at $10,400 for rats, $10,000 for rabbits, $6,500 for mice, $8,600 for hamsters and $55,900 or more for monkeys (44). A recent determination of typical costs for bioassay functions in general was made by Tracor Jitco, Inc., as part of another project, and some relevant excerpts are shown in Table 19. It is emphasized that these determinations are at best approximate, for conditions and test requirements vary enormously. This Table refers to work performed on contract by laboratories specializing in such work. Earlier estimates for work performed at university laboratories specializing in research, shown only for procurement and maintenance of test animals (Table 20) and in fact equally vague and variable, indicate much higher expenses in a university setting than in a contract laboratory. However, this comparison may not be true under all conditions.

If a manufacturer has the primary duty of having teratogenicity tests performed, it is assumed that in most cases the work will be done not 'in-house', but in some laboratory accustomed to doing such work. It is also assumed that the manufacturer will conduct a risk-benefit analysis for his own purposes, on the assumption that he will be held liable for the full costs of damage. The stopper here is the question of insurability and premia.

Table 19

Typical Costs* for Some Carcinogenesis Bioassay Functions Determined by Tracor Jitco, Inc., for the National Cancer Institute in 1976.

Function	Costs
Animal care:	$0.15 per day per animal for an animal population of about 4,000 mice and 4,000 rats.
Necropsy:	$8.50 per adult animal.
Trimming:	$6.50 per animal, 30 organs.
Pathological diagnosis:	$1.60 per tissue.
Histology:	$2.50 per slide, 8 slides per mouse 12 slides per rat or hamster.

*Total costs including labor, materials, overheads, and fees in an average contract laboratory.

Table 20

Laboratory Procurement and Maintenance Costs Determined by the American Association of Medical Colleges for Year Ending June 30, 1973

Animal	Procurement, ready for experiment, $		Husbandry, full costs per day, $
Dogs	45.52-59.65		1.602
Cats	27.50		0.948
Mice	0.50-0.61		0.034
Rats	1.65-2.20		0.170
Guinea pigs	4.40		0.170
Hamsters	1.93		0.034
Rabbits	5.50		0.446
Chickens	2.20		0.258
Calves	110.00		
Swine	33.00	Ungulates	2.022
Frogs	2.20	Amphibians	0.105
Rhesus monkeys	92.62		
Squirrel monkeys	49.00	Primates	1.281

VII. LISTS OF KNOWN TERATOGENS

One comprehensive listing of known teratogens appears to be recognized in the field of teratology: Catalog of Teratogenic Agents by Thomas H. Shepard (130). Published in 1973, this list is computerized with a view to regular updating. Copies have been delivered to the EPA, and the list is not repeated here.

Shepard (130) commented that about 3% of all human newborns have a congenital anomaly requiring medical attention (about 1% being life-threatening) and that over twice this amount are detected with increasing age. About 15% of the total have genetic origins, and less than 3% of the remainder are caused by teratogens. The author listed over 600 teratogens for experimental animals, and stated that only about 20 are known to cause defects in man. He noted that there is a wide gap of knowledge between experimental teratology and the role of external agents in human teratology.

VIII. RECOMMENDATIONS

A. Screening Chemicals Already in the Environment for Teratogenic Effects

1. How chemicals already in the environment would be screened for teratogenic effects is not clear, since there is no readily apparent basis for selecting particular chemicals for screening for potential teratogenicity, from the many thousands presently in the environment, except perhaps a basis of the numbers of people exposed.

2. Nevertheless, it is recommended that a watch be kept on chemicals already in the environment, for the following alterations:

a. Quantity increases that may overstep teratogenicity thresholds, and

b. Interactions, in the environment or in exposed individuals, between chemicals already present and "other chemicals," with teratogenic results. "Other chemicals" might be newly introduced and shown to be nonteratogenic by themselves, or already present and the amount is increased, or introduced as drugs or food chemicals rather than as industrial chemicals, or other perhaps complex alterations of the present state of the environment. Such a watch would be activated by results of population surveillance, mentioned below. Thus if a general or local increase of birth defects, total or specific, were to be revealed by surveillance, inquiry about possible environmental causes would not be limited to newly introduced chemicals but would also cover alterations as outlined above.

3. It is recommended that special exposure groups, which include women, be monitored by surveillance as outlined below for possible teratogenic effects of chemicals in their existing special environments.

4. If, as a result of inquiries in paragraphs 2 and 3 immediately above, suspicion should fall on any chemical already in the general or special environment being surveyed, then that chemical should be screened for potential teratogenicity using criteria as outlined in Subsection B below, and taking into account the circumstances revealed by the surveillance procedures. It is envisaged that the responsibility for the work of screening would be the producer's.

5. Any surveillance of human populations for teratological manifestations can achieve its objectives only if the design is adequate. Following are some recommendations for the design of general surveillance projects and for special surveillance of particular population segments.

a. In practice, areas for surveillance should be selected with geographical or other limits, for example, areas served by selected medical centers already possessing the requisite skills and facilities.

b. All information should be fed into one national collecting center. In Britain all information is fed to the Office of Census and Surveys. In the United States, the Center for Disease Control at Atlanta, Georgia, might be the proper recipient, or FDA or EPA.

c. Within the selected area, reporting should be mandatory.

d. An objective of organized, comprehensive reporting will be to develop baseline criteria, i.e., true background levels of malformations and functional deficits, with the seasonal and other regular fluctations. Then even small irregular 'blips' can be rapidly discerned. Inter-regional comparisons of baseline data will be important.

e. The demography of each reporting area should be recorded and kept up-to-date. All reports of terata should include the mother's age and occupation during pregnancy, place of work, and home address.

f. The environments in the reporting area should be charted (composition of air, water, radiations, and other sources of potential teratogenic exposures) and the map kept up-to-date.

g. In handling and processing the above types of data, the expected timelags relevant to teratogenesis, both before and after exposure of the mother, should not be omitted from the program.

h. The reporting area should be mapped for drug usage: (i) local prescribing patterns by doctors, and (ii) local consumption of over-the-counter drugs by the public, in each case with special reference to pregnant women.

i. A standard system of reporting birth defects and functional deficits is required. Problems in the development of such a system may include: excessive dependence on the knowledge and diagnostic ability of local physicians, excessive need for special procedures, inadequacies in any itemized list of defects at the present time, considerations of privacy for families, irregular performance of procedures by reporting doctors, lack of indicators of many functional deficits and some malformations at time of birth or reporting. Therefore, a special study is recommended to evaluate these and analogous problems and to develop a provisional standardized reporting procedure.

j. Miscarriages, abortions, stillbirths, and deaths in the perinatal period should be included in the mandatory reporting procedures, together with any concomitant malformations. Problems of performance may include: excessive need for autopsy and pathological skills and facilities, resistance by some families in the reporting area, uneven willingness of local doctors to cooperate. The above-mentioned special study recommendation should cover this subsection as well.

k. Notifications should include known pregnancies of employed women, to be reported by personnel departments of the employers.

l. An informative adjunct to human surveillance would be surveillance, at the same time, of hereditary and nonhereditary malformations and functional deficits in the farm and domestic animals in the reporting area. Problems may include: concealments by farmers for commercial reasons or to safeguard pedigree statuses, lack of a mandatory birth-and-death registration system for animals, incomplete involvement of veterinary practitioners in problems of animal health and disease.

m. At the same time note will be taken of current requirements of state governments for notification of birth defects, the extent to which they are fulfilled, the methods used and the problems met. Any state requirements for notification of birth defects in farm animals will also be noted. Where a reporting area includes portions of two or more state jurisdictions, the two sets of requirements may differ. Whether or not this is the case, it is presumed that state governments will be consulted during any surveillance planning and invited to collaborate.

6. Women employed under industrial conditions may be exposed to special teratological risks. Special surveillance is recommended, as follows.

a. Populations should be identified consisting of sufficient numbers of women in one or more areas served by a medical center with the requisite skills and facilities. These may, but need not, be the same as those identified for general surveillance but rather should be selected according to location of types of potential exposure and concentrations of women potentially at risk.

b. Within a population, all women should be subjects, not only pregnant women. Such surveillance is practiced currently in the Soviet Union, and information should be sought through the proper channels.

c. Where the basis of a population is employment in a particular industry, personnel departments of that industry should supply all information required for surveillance purposes, and medical officers already involved in health services to that industry should be involved in the surveillance, ancillary to officers employed or retained as consultants by the regulatory agency.

d. The major routes of exposure for women in these populations are expected to be the skin and the lungs. In addition, food and drink prepared or sold to the population in the areas of potential exposure should also be considered a potential source of exposure, and therefore ingestion may be a potential route.

e. Other recommendations of detail, as mentioned above (Item 5), apply also to special surveillance.

7. All designs for both general and special exposure surveillance procedures should include measures for follow-up in case of positive indications.

B. Prospective Screening of Chemicals Not Yet in the Environment for Potential Teratogenicity

Although doubts will always remain where no evidence of teratogenicity is found, screening for potential teratogenicity is both possible and practical at the present time, and is recommended.

1. Because of the doubts, and because human beings are the best indicators of human teratology, both general and special exposure surveillance procedures are recommended whenever a chemical is permitted into the environment for the first time. Operators of general surveillance will be alerted to watch for appropriate reports and statistical effects at expected times. Special exposure populations will be identified and surveillance prearranged at the time permission is given, and preferably as a condition of that permission.

2. Laboratory screening of "new" chemicals for potential teratogenicity is recommended as follows.

a. Detailed methods must be selected in each case by the manufacturer, in view of his prior and superior knowledge of his product.

b. It is assumed that standard toxicity data will have been determined (acute, subacute, chronic) and that studies on the effects on reproduction of low-level, long-term exposures for several generations will at least have been considered.

c. The teratogenicity study will involve only one generation.

d. It will involve two or more species of rodent or lagomorph, e.g., mouse, rat, hamster, guinea pig, rabbit. If the fetal CNS is a suspected target, a species, such as the guinea pig, having a protracted fetal period comparable to that in humans should be considered. In addition to different species, the comparability of the genetic structure of the test species population to that of the human population must be weighed.

e. In each species selected for testing, one or more identified strains should be used. If results are equivocal, further tests should be performed in additional strains. Debate should be generated on whether inbred or hybrid strains give better information. For example, inbred strains may be better for studies of mechanism, where homogeneity of response is important. In many other studies, outbred strains may give a better representation of the variability to be expected in a heterogeneous population.

f. Each treatment group (test, concurrent controls) should include 20 or more pregnant females, mated in the same laboratory where the assay is to be done, i.e., not transported in the pregnant state. The control group should resemble the test group as closely as possible (e.g., by bodyweight averaging) and should be kept under the same conditions (e.g., temperature, light, humidity, cage-positions, etc.).

g. The LD_{50} of the test substance should be determined with similar animals, e.g., LD_{50} values determined for aged males are irrelevant. If the LD_{50} is impractical, MTD or some other well defined toxic level should be established as a basal criterion.

h. Three dose levels should be employed to begin with. These should be logical fractions of the LD_{50}, e.g., 1/2, 1/4, 1/8; or successive 1/2-dilutions; or log-intervals, or other suitable division, depending on the substance and its characteristics known to the manufacturer. The LD_{50} value should be of pregnant animals, not of the fetus, for the embryotoxic zone may be narrow. If the first three dose-levels used do not result in (i) death of most fetuses at one level, and (ii) survival of most fetuses at another level, then further levels should be administered to other groups until these results are achieved.

i. The typical or expected route of human exposure should be copied as closely as possible. Drinking water, for instance, can be copied exactly, or at least test animals should be dosed by gavage. Aerosols should be given in suitable chambers, although the LD_{50} may be hard to establish this way. The overriding aim is to reproduce in test animals, by whatever route, the tissue levels and persistence of the test chemical and its metabolites that would occur in typical humans under typical exposure conditions. This ideal may be unattainable for many reasons.

j. Dose-equivalents for specific skin surface area (body surface/body-weight) should be adjusted for interspecies comparisons, even if this is not the expected exposure route.

k. The dose schedule should approximate the expected patterns of human exposures (e.g., continuous, intermittent) and conditions of exposure (e.g., sunlight, temperature, humidity), taking into account diurnal and other cycles in the test animal compared with man.

l. The purity of the substance as marketed or released into the environment would be an acceptable criterion, but this should be constant throughout the test series even though it may vary in practice, e.g., between batches. Purity as synthesized would also be acceptable. Alternatively, so would purity in the expected environmental context. The ideal purity and form of the substance is the expected form to which people are likely to be exposed.

m. The test vehicle for the substance to be tested should be inert and should sustain the concentrations and the appropriate physical state (e.g., particulate, non-colloidal) of the substance throughout the experiment. Concurrent controls should be given the vehicle without the substance. The nature of the vehicle will be the manufacturer's responsibility.

n. Pregnant test females should be killed one day or so before expected delivery, and the method should not traumatize the fetus or uterus. If there is particular reason to modify this in order to obtain some meaningful data, some animals may be killed at additional times or by additional or modified methods, e.g., by perfusion-fixation.

o. The abdomen is opened, and the uterus is inspected <u>in situ</u> for total implantation sites including all stages of resorption. The corpora lutea are counted and the total compared with the total observed implantation sites. Preimplantation or early-resorption losses should be estimated either by the foregoing comparison or by other suitable methods. Faulty implantation in mice, rats, and hamsters usually leaves a scar if the embryo survives for at least 24 hours. The fact that accuracy of these estimates is not perfect, and also that it varies by species, does not abolish their usefulness.

p. The uterus is opened, and living fetuses are distinguished from dead, intact fetuses by color, movement and other signs. Macerated fetuses are counted as part resorptions. Early and late resorptions need not be counted separately.

q. The umbilical cord of each living fetus is clamped or cauterized, and the fetus is examined grossly (under magnification if need be) for deviations of proportions, configuration or symmetry, and other gross structural features.

r. Fetuses are blotted and weighed individually.

s. Fetuses are observed for the usual vital signs.

t. The sex of each fetus is recorded.

u. One-half to one-third of fetuses are prepared for skeletal visualization in 95% ethanol as preparation for clearing and staining. The rest are fixed in Bouin's fluid or other nonhardening fixatives that can decalcify the skeleton, as preparation for sectioning or dissection, by standard methods, to reveal any soft-tissue abnormalities. Use of the same fetuses for both skeletal and soft tissue examinations has been reported; the fetuses were fixed in 70% alcohol, dissected, and finally were stained with alizarin (143).

v. If there is reason to suspect that subtle damage of particular systems, e.g., the nervous system, may occur selectively, appropriate procedures should be devised to reveal histological changes. This may need modification of the above procedures or special fixation procedures, e.g., perfusion-fixation. If so, adequate samples should be fixed specially.

w. There is growing concern about possible postnatal functional defects that are not revealed by size, growth, survival, gross anomalies, and other measurements described above. Deviations from normal maturational milestones should be observed, e.g., eyelid opening, hair growth, incisor eruption, freeing of pinna, posture, locomotion. In addition, neuromuscular competence should be tested by appropriate procedures, e.g., swimming, climbing, walking a rod. Because young adults may compensate for such defects, which emerge

when senescent animals are given functional tests, a proportion of test animals may require lifetime observation. This area of concern is considered legitimate, but it evokes problems. Functions that develop in utero in people may develop postnatally in rats. Until suitable animal models or accumulations of background information become available to simulate human time in utero, the best compromise model would involve use of rats because most is known about them. Manufacturers should be prepared to respond to demands for functional tests.

x. The uncertainties in teratogenicity testing require that risk-and-benefit analyses be undertaken. The manufacturer should be responsible for primary data, which can then be checked independently.

IX. POSSIBLE FUTURE METHODOLOGIES

The purpose of this section is to indicate directions of technical development that appear, from data and opinions received during preparation of this report, to be both desirable and currently feasible. Four such items have been identified: postnatal evaluations, shortcuts in current methods of assessing potential teratogenicity, additional species of test animals, and improved monitoring of human populations.

A. Postnatal Evaluations

Much should be known about postnatal physiological and behavioral development of a species before it can be used profitably to assess teratogenic effects that were not readily detected by existing methods at or before birth. Much is known about one species, the rat, because it is used to research the basic processes of postnatal development (137). But its course of postnatal development cannot be compared to that of humans, and so the rat is judged to be less than ideal for the purposes outlined in this survey. Nevertheless, techniques for postnatal evaluation are becoming adequate for screening purposes in teratology, and it would seem desirable and feasible to work up the background knowledge in some more suitable species.

The need for postnatal teratologic evaluation arises mainly because some structures, particularly in the CNS, do not mature fully in human fetuses, in which they may still be liable to certain teratogenic influences even in late pregnancy. As the same structures develop postnatally in many laboratory animals, such animals cannot be used meaningfully for prenatal screening for this teratogenicity. This tends to rule out rats and mice for testing potential late-stage teratogenicity of environmental substances and suggests that the guinea pig (or other suitable species with a protracted fetal period) might be a useful test animal. It seems both desirable and feasible to work up enough background knowledge to qualify the guinea pig for this purpose.

B. Possible Shortcuts in Current Procedures

Shortcuts are desirable and need developing when long lists of chemicals are to be screened for potential teratogenicity in a short time. One

approach would be to consider use of submammalian animals: invertebrates, sub-avian vertebrates, and birds. For example, the metabolism of a chemical by liver microsomes might be studied in vitro, and then submammalian embryos would be treated with it. However, to use this shortcut requires knowledge of basic mechanisms involved, such as the molecular mechanisms by which the chemical damages the tissues that are tested. Such knowledge is often incomplete.

It is not enough to know that two chemicals are structurally related for catalog purposes, for it may be the conformation of 'active' parts of molecules that matters. These may (a) consist of only a minute portion of the molecule, and (b) be similar in substances that would not be classified together by their entire structures. Similar considerations apply to binding-sites of test tissues and organisms. Therefore, testing by chemical analogy demands specific foreknowledge, and there are possible pitfalls. For example, the intact molecule of thalidomide is teratogenic to man during early organogenesis, and its major component, glutamic acid, may be cytotoxic to parts of the CNS in some late-stage animal fetuses. But there is no evidence that glutamic acid is teratogenic during early organogenesis, or thalidomide itself during late pregnancy. Therefore one suspects that these manifestations are chemically unrelated. In summary, research would seem worthwhile and technically feasible in order to identify tissue-chemical interactions that occur both in man and in cell-free systems or simple animal forms lacking maternal-conceptus placental units.

C. Additional Species of Test Animals

Many species that in prospect seem economically feasible have seldom or never been used. The ferret is an example. The principle of species diversification has worked in other fields such as food production, where the gain is protection against epidemics. Suppose, for example, that a test strain were to be standardized and then were to be found sensitive to a mutant pathogen? The background knowledge would become useless. Therefore "tiered" or "multilevel" systems, in which the design of a test involves several species, would seem worth looking into, even though the procedures might become more elaborate. Occasions for use of this sort of procedure would be determined in advance by risk-benefit analysis.

D. Improved Monitoring of Human Populations

Improvements are needed in (a) monitoring, (b) reporting and analyzing (c) dissemination of knowledge, and (d) follow-up action. Better surveillance is needed of the general population and of special groups, mainly women in industry. Interagency coordination seems capable of being improved: agencies within the same government require more coordination, and all sources of information need to be made available nationally and internationally. Some of this may require legislation.

One approach to monitoring would be to select a few potentially high-risk areas or industries in which notification would be compulsory. Data would be collected on all possible anomalies of development, whether these were observed at birth or later, and compiled to facilitate the finding of any connection with any aspect of the environment. The discovery process would involve value judgements. All women in the area would be monitored, and their residences and places of work reported. All doctors, including both obstetricians and other physicians responsible for postnatal health, would be required to participate. Anything that conceivably could be defined as developmental would be noted, and an improved format would be devised. A record linkage system would permit lifetime follow-up of children born during the monitoring. The provisions are of course idealized; to approach them would require research and development of administrative procedures, as well as convincing the public that their privacy is not being invaded.

To facilitate the exchange of information, the March of Dimes National Foundation established the International Clearing House of Birth Defects Monitoring System in 1975. The countries participating in this project are Canada, France, Finland, Hungary, Israel, Norway, Sweden, England and Wales, and the United States. The value of this exchange is, however, limited because of the lack of uniformity in collecting and reporting of data, defining the type of birth defects, etc., and the limited participation in this program of the international community as a whole (94).

X. CONCLUDING REMARKS

If a chemical is found to be teratogenic, should its use ever continue? Under what circumstances? At what levels? These questions of public policy are beyond the scope of this survey, but the answers will largely depend on methods available for testing and in turn will influence the methodologies to be sought.

This survey has revealed that no one routine of testing will reliably indicate whether a chemical is potentially teratogenic for man. The available methods are empirical, and much more fundamental knowledge is needed before teratology testing becomes more rational. For example, extrapolation of data from test animals to man might be closer if extrapolation were possible from one animal species to another. The latter will require extensive comparative teratologic and pharmacologic studies of various classes of chemicals in various species of test animals. A complementary requirement is much improved knowledge of human pharmacology and pharmacologic polymorphisms.

Thus the relationship of screening to research is important, for research is needed to form a basis for organized development of screening methods in teratology.

Even the definition of teratology has caused problems. If teratology results from harmful interference with development of the organism, then to define it by reference to the moment of birth can be misleading. Development is rarely, if ever, complete at birth in any mammalian species, and humans are born at a different stage of development from most test animals. Better criteria of when development is complete might be: (a) the last mitotic cell division in neurons, or (b) the completion of myelination in myelinated neurons. (Sexual maturity is not considered relevant.) This difficulty of definition will be minimized if teratology is considered in proper relationship to other areas of toxicology.

Only by standardizing teratogenic test procedures, by using the best available methodology and well-trained personnel can the existing confusion in comparing teratogenic data, emanating from various laboratories using different animal species, be resolved.

Exposure of the public to teratogenic chemicals appears to be rare, but known occasions have included major disasters. With the continued development and release of potent chemicals into the environment, continued research into screening methods appears to be essential.

Finally, the limitations that exist in predicting human results on the basis of teralogical data obtained from animals led teratologists to oppose legislation that would apply the "Delaney regulation" to teratogens (148). Extension of a Delaney-type clause to include teratogens (such as presently exists for tumor-inducing agents) would have decreed that <u>any</u> agent found to cause malformations at <u>any</u> dose in any experimental animal <u>must</u> be categorized as hazardous and be prohibited from human use. As stated by Karnofsky (73), "the purpose of evaluating drugs for teratogenicity is not to eliminate from use the drugs which show teratogenic potential but rather to assess the hazard their use presents to the human fetus."

REFERENCES

1. Amin-Zaki, L., S. Elhassani, M.A. Majeed, T.W. Clarkson, R.A. Doherty, and M. Greenwood. 1974. Intra-uterine methylmercury poisoning in Iraq. Pediatrics 54(5): 587-595.

2. Anderson, D.H. 1949. Effects of diet during pregnancy upon the incidence of congenital hereditary diaphragmatic hernia in the rat. Am. J. Pathol. 25: 163-185.

3. Anon. 1971. Nutrient Requirements of Poultry, 6th Edition. National Academy of Sciences, National Research Council, Washington, D.C.

4. Anon. 1972. Atlas of Nutritional Data on U.S. and Canadian Feeds. National Academy of Sciences, National Research Council, Washington, D.C.

5. Anon. 1973. The Testing of Chemicals for Carcinogenicity, Mutagenicity, Teratogenicity. Health Protection Branch, Ministry of National Health and Welfare, Canada. pp. vi and 185.

6. Anon. 1974. How Paraquat gets into the lung. New Scientist (December 12, 1974): 797.

7. Armstrong, P.B. 1971. Scanning electron microscope technique for study of internal microanatomy of embryos. Microscope 19: 281.

8. Armstrong, R.C., and J.J. Elias. 1968. Development of embryonic rat eyes in organ culture. I. Effect of glutamine on normal and abnormal development in a chemically defined medium. J. Embryol. Exp. Morphol. 19(3): 397-405.

9. Asling, C.W. 1968. Nutrition and teratogenesis. In: Methods for Teratological Studies in Experimental Animals and Man. H. Nishimura, J.R. Miller, and H. Yasuda, eds. Medical Examination Publishing Co., Flushing, N.Y. pp. 76-91.

10. Becker, B.A. 1974. The statistics of teratology. Teratology 9(3): 261-262.

11. Briggs, J.C., and J.G. Wilson. 1959. Comparison of the teratogenic effects of trypan blue and low temperature in the medakafish (Oryzios latipes). Quart. J. Fla. Acad. Sci. 22: 54-68.

12. Burde, R.M., B. Schainker, and J. Kayes. 1971. Acute effect of oral and subcutaneous administration of monosodium glutamate on the arcuate nucleus of the hypothalamus in mice and rats. Nature 233: 58-60.

13. Cahen, R.L. 1964. Evaluation of the teratogenicity of drugs. Clin. Pharmacol. Therap. 5: 480-514.

14. Cahen, R.L. 1966. Experimental and clinical chemoteratogenesis. Adv. Pharmacol. 4: 264-349.

15. Clayton, F.C., and R.T. Francoeur. 1971. Some teratogenic effects of vinblastine on the external morphology of Drosophilia melanogaster. Oncology 25: 188-192.

16. Clegg, D.J. 1971. Teratology. Ann. Rev. Pharmacol. 11: 409-424.

17. Collaborative Study by 16 Laboratories. 1966. The evaluation of drugs for foetal toxicity and teratogenicity in the rat. Proc. Eur. Soc. Study Drug Tox. 7: 216-228.

18. Chaube, S., and M.L. Murphy. 1969. The teratogenic effects of recent drugs active in cancer chemotherapy. In: Advances in Teratology. D.H.M. Woollam, ed. Logos Press Ltd., London. Vol 3. p. 205.

19. Collins, T.F.X., and C.H. Williams. 1971. Teratogenic studies with 2,4,5-T and 2,4-D in the hamster. Bull. Environ. Contam. Toxicol. 6(6): 559-567.

20. Cooke, A.S. 1972. The effects of DDT, Dieldrin, and 2,4-D on amphibian spawn and tadpoles. Environ. Pollut. 3(1): 51-68.

21. Cooke, A.S. 1973. Effect of DDT, when used as a mosquito larvacide, on tadpoles of the frog Rana temporaria. Environ. Pollut. 5: 259-273.

22. Critchfield, C., and J.C. Daniel, Jr. 1965. Teratogenic effects of trypan blue on Coturnix quail when injected into the mother. Growth 29(3): 301-309.

23. Dawson, A.B. 1926. A note on the staining of the skeleton of cleared specimens with alizarin red S. Stain Technol. 1: 123-124.

24. Delahunt, C.S. 1970. Detection of teratogenic actions. In: Methods in Toxicology. G.E. Paget, ed. Blackwell Scientific Publications, Oxford, England. pp. 132-157.

25. Domac-Tesar, B., and Z. Frank. 1963. Effects of local application hypothermia during embryonic development of the white rat (Abstract). Acta Anat. 54: 353.

26. Dostal, M., and R. Jelinek. 1971. Induction of cleft palate in rats with intra-amniotic corticoids. Nature 230: 464.

27. Duby, R.T., and H.F. Travis. 1971. Influence of dienestrol diacetate on reproductive performance of female mink (Mustela vision). Am. J. Vet. Res. 32(10): 1599-1602.

28. Evans, M.A., M.W. Stevens, B. Mantilla-Plata, and R.D. Harbison. 1975. Drugs of abuse: teratogenic and mutagenic considerations. Addictive Diseases. 2: 45-61.

29. Faherty, J.F. 1972. Surface staining of 1 mm (Wilson) slices of fetuses for internal visceral examination. Stain Technol. 47: 53-58.

30. Feldmann, B.M., L.M. Dudman, and M.S. Redfearn. 1971. Canine, feline and lab primate dry feeds: guidelines to evaluation. Lab. Anim. Sci. 21(6): 862-864.

31. Ferm, V.H. 1965. The rapid detection of teratogenic activity. Lab. Invest. 14: 1500-1505.

32. Ferm, V.H. 1967. Potentiation of the teratogenic effect of vitamin A with exposure to low environmental temperature. Life Sci. 6: 493-497.

33. Ferm, V.H., and A.R. Beaudoin. 1965. Studies on the effect of trypan blue in the pregnant armadillo, Dasypus novemcinctus. Anat. Rec. 151(4): 571-578.

34. Ferngren, H., and U. Forsberg. 1971. The correlation of adverse effects in man with observations in animals. Evaluation in animals of teratogenic effects of drugs submitted to the Swedish Drug Control 1963-1968. Proc. Eur. Soc. Study Drug Tox. 12: 347-351.

35. Fitzhugh, O.G. 1968. Reproduction tests. In: Modern Trends in Toxicology. E. Boyland, and R. Groulding, eds. Butterworth, London. vol. 1. pp. 75-85.

36. Flynt, J.W., Jr. 1974. Trends in surveillance of congenital malformations. In: Congenital Defects. D.T. Janerich, R.G. Skalko, and I.H. Porter, eds. Academic Press, New York. pp. 119-127.

37. Flynt, J.W., Jr., A.J. Ebbin, G.P. Oakley, Jr., A Falek, and C.W. Heath. 1971. Metropolitan Atlanta congenital defects program. In: Monitoring, Birth Defects and Environment. E.B. Hook, D.T. Janerich, and I.H. Porter, eds. Academic Press, New York. pp. 155-158.

38. Forsberg, U., and C.A. Grant. 1971. The correlation of adverse effects in man with observations in animals. Current practices of pharmaceutical companies for detecting and reporting foetal anomalies obtained in animal tests. Proc. Eur. Soc. Study Drug Tox. 12: 352-356.

39. Fraser, F.C. 1964. Experimental teratogenesis in relation to congenital malformations in man. In: Second International Conference on Congenital Malformations. Lippincott, Philadelphia. pp. 277-287.

40. Fraser, F.C., and V.A. McKusick, eds. 1970. Congenital Malformations. Excerpta Medica, North Holland Publishing Co., Amsterdam.

41. Freireich, E.F., E. Frei, III, and M. Karon. 1962. Methylglyoxal bis-(guanylhydrazone): a new agent active against acute myelocytic leukemia. Cancer Chemother. Rep. 16: 183-186.

42. Friedman, L., C.M. Kunin, N. Nelson, J.L. Whittenberger, and J.G. Wilson. 1970. Food and Drug Administration Advisory Committee on Protocols for Safety Evaluations: Panel on Reproduction report on reproduction studies in the safety evaluation of food additives and pesticide residues. Toxicol. Appl. Pharmacol. 16: 264-196.

43. Gebhardt, D.O.E., and J. Faber. 1966. The influence of aminopterin on limb regeneration in Ambystoma mexicanum. J. Embryol. Exp. Morphol. 16(1): 143-158.

44. Gehring, P.G., V.K. Rowe, and S.B. McCollister. 1973. Toxicology: Cost/time. Food Cosmet. Toxicol. 11: 1097-1110.

45. Gilbert, E.F., H.C. Pitot, H.J. Bruyere, Jr., and A.L. Cheung. 1973. Teratogenic effects of 5-bromodeoxyuridine on the external morphology of Drosophila melanogaster. Teratology 7: 205-208.

46. Giroud, A. 1970. Nutrition of the Embryo. Thomas, Springfield, Ill.

47. Gofmekler, V.A. 1969. Methods for studying the effects of chemical substances on embryos. Hyg. Sanit. 34: 198-202.

48. Goldstein. A., L. Aronow, and S.M. Kalman. 1969. Principles of Drug Action. Harper and Row,New York. pp. 180.

49. Goodhart, R.S. 1964. Criteria of an adequate diet. In: Modern Nutrition in Health and Disease. M.G. Wohl, and R.S. Goodhart, eds. Lea & Febiger, Philadelphia. pp. 617-632.

50. Greenhouse, G., and M. Hamburgh. 1968. Analysis of trypan blue induced teratogenesis in Rana pipiens embroys. Teratology 1: 61-74.

51. The effect of serotonin on the offspring and the fetal-maternal relationship in rats. Arzneim.-Forsch. 25(4): 626-632.

52. Hagstrom, B.E., and S. Lonning. 1973. The sea urchin egg as a testing object in toxicology. Acta Pharm. Toxicol. 32(Suppl. 1): 1-49.

53. Hamburgh, M., L.A. Mendoza, M. Rader, A. Lang, H. Silverstein, and K. Hoffman. 1974. Malformations induced in offspring of crowded and parabiotically-stressed mice. Teratology 10(1): 31-37.

54. Hanley, T., V. Udall, and M. Weatherall. 1970. An industrial view of current practice in predicting drug toxicity. Br. Med. Bull. 26: 203-207.

55. Harbison, R.D., and B.A. Becker. 1971. Effects of phenobarbital and SKF 525A pretreatment on phenylhydantoin disposition in pregnant mice. Toxicol. Appl. Pharmacol. 20: 571-581.

56. Hart, N.H., and M. Greene. 1971. LSD: teratogenic action in chick blastoderms. Proc. Soc. Exp. Biol. Med. 137(2): 371-373.

57. Haseman, J.K., and M.D. Hogan. 1975. Selection of experimental unit in teratology studies. Teratology 12: 165-171.

58. Hendrickx, A.G., L.R. Axelrod, and L.D. Clayborn. 1966. Thalidomide syndrome in baboons. Nature 210: 958-959.

59. Hill, G.B., C.C. Spicer, and J.A.C. Weatherall. 168. The computer surveillance of congenital malformations. Br. Med. Bull. 24: 215-218.

60. Horar, R.M., and A.J. Salem. 1961. Time of teratogenic action of trypan blue in guinea pigs. Anat. Rec. 141: 173-181.

61. Humphrey, R.R. 1967. Genetic and experimental studies on a lethal trait ("short toes") in the Mexican axolotl (Ambystoma mexicanum). J. Exp. Zool. 164: 281-296.

62. Hurley, L.S. 1965. Nutritional Factors in Teratogenesis. Supplement to the Teratology Workshop Manual, Pharmaceutical Manufacturers Assn., Washington, D.C. pp. 114-120.

63. Hurley, L.S., and H. Swenerton. 1966. Congenital malformations resulting from zinc deficiency in rats. Proc. Soc. Exp. Biol. Med. 123: 692-696.

64. Ingalls, T.H., and F.J. Curley, 1957. Principles governing the genesis of congenital malformations induced in mice by anoxia. New Engl. J. Med. 257: 1121-1127.

65. Jensh, R.P., R.L. Brent, and M. Barr. 1970. The litter effect as a variable in teratologic studies of the albino rat. Am. J. Anat. 128: 185-192.

66. Jones, D.C. 1969. Persistent and late effects of whole-body irradiation in juvenile male rats. In: Radiation Biology of the Fetal and Juvenile Mammal. M.R. Sikov and D.D. Mahlum, eds. U.S. Atomic Energy Commission, Washington, D.C., pp. 439-447.

67. Jones, R.W., and M.N. Huffman. 1957. Fish embryos as bioassay material in testing chemicals for effect on cell division and differentiation. Trans. Am. Microscop. Soc. 76: 177-183.

68. Kalter, H. 1977. The structure and uses of genetically homogenous lines of animals. In: Handbook of Teratology. J. C. Wilson and F.C. Fraser, eds. Plenum Press, New York.

69. Kalter, H. 1968. Teratology of the Central Nervous System. University of Chicago Press, Chicago, Ill.

70. Kalter, H. 1975. Choice of the number of sampling units in teratology. Teratology 9(3): 257-258.

71. Kalter, H. 1975. Personal communication.

72. Kalter, H., and J. Warkany. 1959. Experimental production of congenital malformations by metabolic procedures. Physiol. Rev. 39: 69-115.

73. Karnofsky, D.A. 1965. Drugs as teratogens in man. Ann. Rev. Pharmacol. 5: 447-472.

74. Karnofsky, D.A. 1965. The chick embryo in drug screening: survey of teratological effects observed in the 4-day chick embryo. In: Teratology: Principles and Techniques. J.G. Wilson, and J. Warkany, eds. University of Chicago Press. pp. 194-213.

75. Karnofsky, D.A., and E.B. Simmel. 1963. Effects of growth-inhibiting chemicals on the sand-dollar embryo (Echinarachinus parma). Prog. Exp. Tumor Res. 3:2 254-295.

76. Karnofsky, D.A., and R.S. Basch. 1960. Effects of 5-fluorodeoxyuridine and related halogenated pyrimidines on the sand-dollar embryo. J. Biophys. Biochem. Cytol. 7(1): 61-71.

77. Keeler, R.F. 1973. Lupin alkaloids from teratogenic and nonteratogenic lupins. I. Correlation of crooked calf disease incidence with alkaloid distribution determined by gas chromotography. Teratology 7: 23-30.

78. Kelsey, F.O. 1974. Present guidelines for teratogenic studies in experimental animals. In: Congenital Defects. D.T. Janerich, R.G. Skalko, and I.H. Porter, eds. Academic Press, New York. pp. 195-204.

79. Kenaga, E.E. 1976. Use of biological tests for evaluation of pesticides. Pure Appl. Chem. 42: 285-299.

80. Kennedy, L.A., and T.V.N. Persaud. 1976. Experimental amniocentesis and teratogenesis. Anat. Anz. Bd. 140: 267-276.

81. Khera, K.S. 1973. Teratogenic effects of methylmercury in the cat: a note on the use of this species as a model for teratogenicity studies. Teratology. 8(3): 293-303.

82. Kirkpatrick, J.F. 1973. Radiation induced abnormalities in early in vitro mouse embryos. Anat. Rec. 176: 397-404.

83. Klein, N.W. 1965. Teratological studies with explanted chick embryos. In: Teratology: Principles and Techniques. J.G. Wilson, and J. Warkany, eds. University of Chicago Press. pp. 131-141.

84. Krowke, R., and D. Neubert. 1977. Embryonic intermediary metabolism under normal and pathological conditions. In: Handbook of Teratology. J.G. Wilson and F.C. Fraser, eds. Plenum Press, N.Y. Vol. 2. pp. 117-151.

85. Kupper, L.L., and J.K. Haseman. 1977. On the use of a correlated binomial model for the analysis of certain toxicological experiments. "Tentatively Accepted" for publication in Biometrics.

86. Landauer, W. 1957. Phenocopies and genotype, with special reference to sporadically-occurring developmental variants. Am. Natur. 91: 79-90.

87. Langman, J., W. Webster, and P. Rodier. 1975. Morphological and behavioral abnormalities caused by insults to the CNS in the prenatal period. In: Teratology: Trends and Applications. C.L. Berry, and D.E. Poswillo, eds. Springer-Verlag, Berlin, Heidelberg, New York. pp. 183-199.

88. Lehman, J., D. Arnold, W. Fasset, W. Horace, H. Guarde, E. Stockinger, and J.W. Zapp. 1964. Principles and Procedures for Evaluating the Toxicity of Household Substances. National Academy of Sciences, Washington, D.C.

89. Lemkey-Johnson, N., and W.A. Reynolds. 1972. Incidence and extent of brain lesions in mice followoing ingestion of monosodium glutamate. Anat. Rec. 172: 354.

90. Lin, F.O., and J.K. Haseman, 1976. A modified Jonckheere test against ordered alternatives when ties are present at a single extreme value. Biom. Z. Bd. 8: 623-631.

91. Lustig, E.S. 1971. Teratogenic effects induced in tail of Bufo arenarum tadpoles following treatment wth carcinogens. Experientia 27: 555-556.

92. Milham, S. 1963. Malformation surveillance in New York State. New York State Med. J. 63: 2823-2824.

93. Miller, H.C., and K. Hassanein. 1971. Diagnosis of impaired fetal growth in newborn infants. Pediatrics 48: 511-522.

94. Miller, J.R. 1976. Birth defect monitoring system: An overview. Teratology 14: 235.

95. Miller, R.W. 1971. Monitoring congenital malformations. In: Monitoring, Birth Defects and Environment. E.B. Hook, D.T. Janerich, and I.H. Porter, eds. Academic Press, New York. pp. 97-111.

96. Mitchell, J.T., and C.L. Yntema. 1973. Teratogenic effect of malathion and captan in the embryo of the common snapping turtle Chelydra serpentina. Anat. Rec. 175(2): 390.

97. Mnatsakanyan, A.V., U.G. Pogosyna, K.Kh. Akopyan, V.A. Gofmekler, A.O. Avoyan, and M.S. Andikyan. 1971. Certain aspects of the embryotoxic effect of chloroprene in a field experiment. Hyg. Sanit. 36(7): 140-141.

98. Moni, I.W. 1965. Comparative development of rat, chick, and human embryos. In: Supplement to Teratology Workshop Manual. PMA, Berkley. pp. 154.

99. Mould, G.P., S.H. Curry, and F. Beck. 1973. The ferret, a useful model for teratogenic study. Naunyn-Schmiedeberg Arch. Pharmacol. 279(24): R-18.

100. Mussman, H.C. 1975. Drug and chemical residues in domestic animals. Fed. Proc. 34(2): 197-201.

101. March of Dimes National Foundation. 1975. Facts. National Foundation. New York.

102. Neuweiler, W., and R.H.H. Richter. 1961. Etiology of gross malformations. Schweiz. Med. Wochenschr. 91: 359-363.

103. New, D.A.T. 1955. A new technique for the cultivation of the chick embryo in vitro. J. Embryol. Exp. Morphol. 3: 326-331.

104. New, D.A.T. 1966. Culture of Vertebrate Embryos. Academic Press, New York, N.Y.

105. New, D.A.T., and K.F. Stein. 1963. Cultivation of mouse embryos in vitro. Nature 199: 297-299.

106. New, D.A.T., and R.L. Brent. 1972. Effect of yolk-sac antibody on rat embryos grown in culture. J. Embryol. Exp. Morphol. 27: 543-553.

107. Nielsen, N.O. 1969. Teratogenic effects of hypothermia in chick embryos. In: Teratology. A. Bertelli, and L. Donati, eds. Excerpta Med. Found., Amsterdam. pp. 102-107.

108. Nishimura, H., and T. Tanimura. 1976. Clinical Aspects of the Teratogenicity of Drugs. Excerpta Medica, Amsterdam (American Elsevier Publishing Co. Inc., New York). pp. 271-298.

109. Nishimura, H., and Y. Yasuda. 1969. Culture of human and mouse limb buds as a tool for testing the embryotoxicity of drugs. In: Methods for Teratological Studies in Experimental Animals and Man. Medical Examination Publishing Co., Flushing, N.Y. pp. 272-273.

110. Nishimura, H., J.R. Miller, and M. Yasuda, eds. 1969. Methods for Teratological Studies in Experimental Animals and Man. Medical Examination Publishing Co., Flushing, N.Y.

111. Noback, G.J. 1916-1917. The use of the Van Wijhe method for the staining of the cartilaginous skeleton. Anat. Rec. 11: 292-294.

112. Nomura, T. 1968. Management of animals for use in teratological experiments. In: Methods for Teratological Studies in Experimental Animals and Man. Medical Examination Publishing Co., Flushing, N.Y. pp. 3-15.

113. Olney, J.W. 1971. Glutamate-induced neuron necrosis in the infant mouse hypothalamus. J. Neuropathol. Exp. Neurol. 30(1): 75-90.

114. Olney, J W., and O-L. Ho. 1970. Brain damage in infant mice following oral intake of glutamate, asparatate, or cysteine. Nature 227: 609-611.

115. Palmer, A.K. 1969. The relationship between screening tests for drug safety and other teratological investigations. In: Teratology. A. Bertelli, and L. Donati, eds. Excerpta Medica Foundation, Amsterdam. pp. 55-72.

116. Paschke, M., and F. Schwanitz. 1968. On the problems of pretesting (preliminary testing) drugs with the acid of lower and higher plants. I. Effect of phthalylglutamic acid imide (thalidomide) and other sedatives and hypnotics on rate of cell division, shape, and sexuality of some Desmidiaceae. Arzn.-Forsch. 18(10): 1338-1341.

117. Peck, H.M. 1963. Preclinical evaluation of drugs for evidence of teratogenic activity. J. Pharm. Sci. 52: 1115-1120.

118. Persaud, T.V.N. 1969. Fetal toxicity of leucine in the rat. West Indian Med. J. 18(1): 34-39.

119. Potts, E.L. 1962. Pathology of the Fetus and the Infant. Year Book Medical Publishers, Inc., Chicago, Ill.

120. Rieselbach, R.E., E.E. Morse, D.P. Rall, E. Frei, III, and E.J. Freireich. 1962. Treatment of meningeal leukemia with intrathecal aminopterin. Cancer Chemother. Rep. 16: 191-196.

121. Robkin, M.A., T.H. Shepard, and T. Tanimura. 1972. A new in vitro culture technique for rat embryos. Teratology 5: 367-376.

122. Robson, J.M. 1970. Testing drugs for teratogenicity and their effect on fertility. Br. Med. Bull. 26: 212-216.

123. Rodier, P., W. Webster, and J. Langman. Morphological and behavioral consequences of chemically-induced lesions of the CNS. In: Proc. 7th Gatlinburg Conf. on Research and Theory in Mental Retardation. N. Ellis, ed. Lawrence Erlbaum Assoc., Potomac, Md.

124. Rosenkrantz, J.G., W.W. Frost, and J.V. Krupp. Unpublished. Effects of trypan blue on the Pitman-Moore miniature pig.

125. Runner, N.M. 1971. The chick embryo as a test system for teratogenic drugs. In: Proc. 3rd Teratology Workshop. Boulder, Colorado.

126. Runner, N.M. 1959. Inheritance of susceptibility to congenital deformity. Metabolic clues provided by experiments with teratologenic agents. Pediatrics 23: 245-251.

127. Saxen, L., M. Karkinen-Joaskelainen, and I. Saxen. 1976. Organ culture in teratology. Curr. Top. Pathol. 62: 123-143.

128. Schardein, J.L. 1976. Drugs as Teratogens. CRC Press, Cleveland, Ohio. p. 22-25.

129. Schnell, V., and J.W. Newberne. 1970. Accelerated clearing and staining of teratologic specimens by heat and light. Teratology 3: 345-347.

130. Shepard, T.H. 1973. Catalog of Teratogenic Agents. Johns Hopkins University Press, Baltimore, Maryland.

131. Shimizu, K., A. Mizutani, and M. Inoup. 1973. Electron microscopic studies on the hypothalamic lesions in the mouse fetus caused by monosodium glutamate (Abstract). Teratology 8: 105.

132. Sit, K.H., an R. Kanagasuntheram. 1972. A structural analysis of congenital limb deformities in experimental hyperthyroid tadpoles. J. Embryol. Exp. Morphol. 28: 223-234.

133. Sladecek, F., and A. Romanovsky. 1960. The effect of cold and heat shock on the eggs of axolotl, recorded to early larval stages. Folia Biol. (Prague) 6: 42-49.

134. Smithberg, M. 1962. Teratogenic effects of tolbutamide on early development of the fish, Oryzias latipes. Am. J. Anat. 111(2): 205-213.

135. Somers, J., E.T. Moran, Jr., and B.S. Reinhart. 1974. Effect of external application of pesticides to the fertile egg on hatching success and early chick performance. 2. Commercial-herbicide mixtures of 2,4-D with picloram or 2,4,5-T using the pheasant. Bull. Environ. Contam. Toxicol. 11(4): 339-342.

136. Spyker, J.M. 1975. Assessing the impact of low-level chemicals on development: behavioral and latent effects. Fed. Proc. 34(9): 1835-1844.

137. Spyker, J.M. 1975. Behavioral teratology and toxicology. Environ. Sci. Res. 5: 311-349.

138. Staples, R.E., and J.K. Haseman. 1974. Selection of appropriate experimental units in teratology. Teratology. 9(3): 259-260.

139. Steffek, A.J., and A.C. Verrusio. 1972. Experimentally induced oral-facial malformations in the ferret (Mustela putorius furo). Teratology 5: 268.

140. Stewart, R.J.C., R.F. Preece, and H.G. Sheppard. 1975. Twelve generations of marginal protein deficiency. Br. J. Nutr. 33: 233-253.

141. Streeter, G.L. 1933. A human embryo of the presomite period. Carnegie Contributions. Embryol. 9: 389.

142. Stula, E.F., and W.C. Kraus. 1977. Embryotoxicity in rats and rabbits from cutaneous application of amide-type solvents and substituted ureas. Toxicol. Appl. Pharmacol. 41: 35-55.

143. Sullivan, F.M., and P.R. McElhatton. 1975. Teratogenic activity of the antiepileptic drugs phenobarbital, phenytoin and primidone in mice. Toxicol. Appl. Pharmacol. 34: 271-282.

144. Sullivan, F.M. 1974. Animal tests to screen for human teratogens. Pediatrics 53: 822-823.

145. Szabo, K.T., and A.K. Palmer. 1974. Teratological evaluation of ophthalmic preparations by the ocular, dermal, and intraperitoneal routes of administration in the rabbit. Teratology 9(3): A38-A39.

146. Takagaki, S., and H. Kanamori. 1957. Developmental anomalies of embryos of Bufo vulgaris induced by 2,3-dimercaptopropanol (BAL). Acta Anat. Nippon. 32(3): 326.

147. Tencer, R. 1961. The effects of 5-fluorodeoxyuridine on amphibian embryos. Exp. Cell Res. 23: 418-419.

148. Teratology Society. 1974. Teratogens and Delaney clause. Teratology 10: 1.

149. Tinker, J. 1974. Vapona: How much is too much? New Scientist (December 12, 1974).

150. Tracor Jitco, Inc. 1974. GRAS Monograph Series: Glutamates. National Technical Information Service, Springfield, Va. No. PB.229856/AS.

151. Tuchmann-Duplessis, H. 1966. Methods for evaluating teratogenic properties of new drugs. In: Methods in Drug Evaluation. P. Mantegazza, and F. Piccini, eds. North Holland Publishing Co., Amsterdam. pp. 11-31.

152. Tuchmann-Duplessis, H. 1970. The problems of species difference and statistics. Proc. Eur. Soc. Study of Drug Toxicity. 11: 33-49.

153. Tuchmann-Duplessis, H. 1972. Teratogenic drug screening. Present procedures and requirements. Teratology 5: 271-285.

154. Turbow, M.M., and J.G. Chamberlain. 1968. Direct effects of 6-aminonicotinamide on the developing rat embryo in vitro and in vivo. Teratology 1: 103-108.

155. U.S. Dept. of Health, Education and Welfare. 1966. Guidelines for Reproduction Studies for Safety Evaluation of Drugs for Human Use. Food and Drug Administration, Washington, D.C.

156. Wainwright, S.D., L.K. Wainwright, and I.H. Fraser. 1973. Hemoglobin formation in the explanted chick blastodisc. A model system for evaluation of embryotoxic agents as illustrated by effects of vinblastine, 1,1,1-trichloro-2,2-bis-(p-chlorophenyl)-ethane (DDT), and sodium diethyl barbiturate. Toxicol. Appl. Pharmacol. 25: 123-129.

157. Warkany, J. 1969. Experimental production of mammalian limb malformations. In: Limb Development and Deformity. Chester A. Swinyard, ed. Thomas, Springfield, Ill. pp. 140-160.

158. Warkany, J. 1971. Congenital Malformations. Year Book Medical Publishers, Inc., Chicago, Ill.

159. Warkany, J. 1973. Congenital malformations of the brain produced by zinc deficiencies in rats. Am. J. Ment. Defic. 77: 645-653.

160. Weatherall, J.A.C. 1969. An assessment of the efficiency of notification of congenital malformations. Med. Off. 121(6): 65-68.

161. Weatherall, J.A.C. 1970. The detection of incidence of congenital malformations in the community. Proc. Roy Soc. Med. 62(12): 1251-1252.

162. Weil, C.S. 1970. Selection of the valid number of sampling units and a consideration of their combination in toxicological studies involving reproduction, teratogenesis or carcinogenesis. Fd. Cosmet. Toxicol. 8: 177-182.

163. Wilson, J.G. 1975. Critique of current methods for teratogenic testing in animals and suggestions for their improvement. In: Methods for Detection of Environmental Agents That Produce Congenital Defects. T.H. Shepard, J.R. Miller, and M. Marois, eds. North Holland Publishing Co., Amsterdam. pp. 29-49.

164. Wilson, J.G. 1965. Methods for administering agents and detecting malformations in experimental animals. In: Teratology: Principles and Techniques. J.G. Wilson and J. Warkany, eds. University of Chicago Press, Chicago, Ill.

165. Wilson, J.G. 1973. Environment and Birth Defects. Academic Press, New York.

166. Wilson, J.G. 1974. Factors determining the teratogenicity of drugs. Ann. Rev. Pharmacol. 14: 205-217.

167. Wilson, J.G. 1975. Personal communication.

168. Wilson, J.G. 1975. Reproduction and teratogenesis: Current methods and suggested improvements. J. Assoc. Off. Anal. Chem. 58: 657-667.

169. Wilson, J.G., and J. Warkany, eds. 1965. Teratology: Principles and Techniques. University of Chicago Press, Chicago, Ill.; p. 279.

170. Wilson, J.G., T.H. Maren, K. Takano, and A. Ellison. 1068. Teratogenic action of carbonic anhydrase inhibitors in the rat. Teratology 1(1): 51-60.

171. Winick, M. 1974. Cellular growth during normal and abnormal development of the brain. In: Biochemistry of the Developing Brain. W. Himwich, ed. Marcel Dekker Inc., New York. Vol 2. p. 199.

172. Witschi, E. 1956. Development of the Vertebrates. W.B. Saunders, Philadelphia.

173. World Health Organization. 1967. Principles for the Testing of Drugs for Teratogenicity. WHO Tech. Rep. Ser. No. 364.

174. Yoshida, Y., T. Ohashi, K. Takano, N. Maekawa, N. Toyoda, R. Hujimoto, K. Yamamoto, S. Iwata, H. Ito, and Y. Mikami. 1957. Visceral malformations produced by vital dyes, Janus green, trypan blue and methylene blue in amphibian embryos. Mie Med. J. 7(2): 273-285.

175. Yoshihara, H. 1966. The effect of 5-flurururacil on development of tail vertebrae of mice in organ culture. Anat. Rec. 154: 445.

INDEX

	Pages
Azo dyes	9
Chemicals	1-3, 6, 9, 10, 19-36, 51, 53, 54, 65, 69, 75, 77, 78
Cleft Lip	47
Cleft palate	18, 46
Cortisone	49
Death Certificate, infant	58, 60
Death, intrauterine	2, 46, 48, 54
Defects, brain	18
Defects, central nervous system	6, 26, 34, 40, 69, 74, 75
Defects, definitions	2, 11, 40-41, 47-48, 58
Defects, digits	47
Defects, eye	34
Defects, functional	2, 37, 38, 41, 67, 72, 74
Defects, skeletal	18, 34, 38, 41, 47, 54, 72
Defects, soft-tissue	18, 34, 38, 41-45, 54, 72-73
Development, postnatal	6-7, 11, 27, 32, 41-45, 74, 77
Development, structural	4, 12, 16, 26, 27, 34, 41, 43, 44
Dietary factors	2, 6, 10, 22-23, 26, 35-38, 46, 54, 60 65, 68
Dosage	6-9, 19-26, 34-36, 47, 53-55, 70, 71
Dose vehicle	8-9, 19, 26-33, 42, 55, 71
Drugs	3, 38, 49, 53, 55, 66
Economics	16, 58, 61-63, 67, 73, 75
Embryotoxicity	8, 21-23, 33, 48-49, 53, 55

INDEX (Cont'd)

	Pages
Environmental factors	1-6, 8-9, 12, 20, 22-25, 34-38, 50-51, 58, 64-66, 68, 70-71, 76-77
Examination, biochemical	45, 75
Examination, visual	7, 11, 38-45, 54, 62, 71-72
Examination, histological	7, 11, 38, 41-45, 62, 72
Exposure	2-6, 8-9, 19-20, 22-24, 55, 59-60, 64-66, 68-71, 70-76
Extrapolability	3-6, 12, 19-20, 22-23, 47-48, 50-51, 64, 70, 72, 74-75, 77
Fetus	6, 27-29, 38-46, 54-55, 69-72, 74
Galactoflavin	13, 18
Gestation	4, 28, 39, 42, 43, 53-55
Glutamic acid	75
Growth retardation	2, 38-39, 46-47
Guidelines	26, 33, 53-56
Herbicides	9
Litter effect	49-50
Malathion	51
Malformations	2, 7-11, 13, 18, 29, 37, 39, 40-41, 45, 47-48, 55, 58, 66-67
Metabolic tolerance	8, 24
Metabolism	4-6, 10, 12, 24, 26, 30, 50-51, 54, 75
Necropsy	7, 39-41, 57, 67, 71-72
Organogenesis	4-6, 10, 25, 27, 31, 35, 48, 53, 55, 75
Paraquat	9
Pesticides	60

	Pages
Placenta	6, 13, 23, 27, 40, 75
Radiation	2, 46, 66
Reporting, centers for	58-60, 66-68, 76
Reporting, defects in farm and domestic animals	60, 67
Reporting, monitoring defects	57-60, 66-68, 76
Risk-benefit analysis	12, 13, 49, 58, 61-63, 73, 75-76
Resorptions	71
Route of exposure or administration	6-7, 22-24, 35, 55, 70
Screening, chemicals	11, 19, 21-22, 24-25, 56, 61, 65-73, 74, 77
Spina bifida	34
Stillbirth	2, 53, 58-60, 67
Strontium-90	9
Surveillance	3, 4, 57-60, 65-68, 69, 76
Teratogen	2, 9, 35, 37-38, 47, 49, 59, 64, 77
Teratogenic ratio	8
Test animals	4-6, 8-9, 12-16, 28, 33-36, 37-38, 52 53-60, 64, 69-72, 74-75, 77
Testing, controls	6, 8-9, 26-36, 45-49, 54, 69-71
Testing, errors	7, 50, 67
Testing, methods	1-9, 12-45, 49-56, 65, 69-72, 74-75, 77
Testing, multigeneration	22
Testing, multilevel	12, 15, 75
Testing, results	1-3, 8-10, 38-40, 45-51

INDEX (Cont'd)

	Pages
Testing, variables	4-6, 11, 13, 16, 21, 33-38, 50
Thalidomide	4, 16, 17-18, 48, 75
Trypan blue	9, 49